Fuzzy quantifiers: a linguistic technique for data fusion

Alois Knoll and Ingo Glöckner

Universität Bielefeld, Technische Fakultät
Arbeitsgruppe Technische Informatik
Postfach 10 01 31, D-33501 Bielefeld
knoll,ingo@techfak.uni-bielefeld.de

Abstract Fuzzy quantifiers like “few”, “almost”, “about” and many others abound in natural language. They are used by humans for describing uncertain facts, quantitative relations and processes. An adequate contradiction-free computer-operational implementation of these quantifiers would provide a class of powerful yet human-understandable operators both for aggregation and fusion of data but also for steering the fusion process on a higher level through a safe transfer of expert-knowledge expressed in natural language. In this chapter we show by a number of examples of image data that the traditional theories of fuzzy quantification (Sigma-count, FE-count, FG-count and OWA-approach) are linguistically inconsistent and produce implausible results in many common and relevant situations. To overcome the deficiencies of these approaches, we developed a new theory of fuzzy quantification, DFS, that rests on the foundation of the theory of generalised quantifiers TGQ. It provides a linguistically sound basis for the most important case of multi-place quantification with proportional quantifiers. Its axiomatic basis guarantees compliance with linguistic adequacy considerations. The underlying models generalize the basic FG-count approach/Sugeno integral and the basic OWA approach/Choquet integral. We have also developed an efficient implementation based on histogram computations. At the end of the chapter the power of the theory and its implementation are illustrated by image data examples.

1 Introduction and motivation

Fuzzy quantifiers (like **many**, **few**. . .) are an important research subject not only due to their abundance in natural language, but also because an adequate computer-operational implementation of these quantifiers would provide a class of powerful yet human-understandable operators for information aggregation and data fusion. Recognising their value as genuine operators for the aggregation over *sets* of gradual evaluations (i.e. irreducible to pairwise combination of results), the use of fuzzy quantifiers has been suggested in the literature for various purposes of aggregation and data fusion and in a variety of applications including multi-criteria decision making, fuzzy databases and information retrieval, fuzzy expert systems, and others. These attempts are limited in use due to their lack of theoretical foundation, which can result in counter-intuitive behaviour as reported e.g. by Ralescu (1986), Ralescu (1995), Yager (1993) and Glöckner (1999).

The present work originated from the need for an adequate interpretation of natural language statements with respect to sets of data that have to be aggregated and fused under linguistic criteria. This need arose in the construction of a large scale database system of multimedia weather documents, cf. Knoll *et al.* (1998), Glöckner and Knoll (1999b). It accepts queries that result in an aggregation of information and queries that require the fusion of different forms of data. The *natural language* (NL) *queries*, e.g. “*Show me recent cloudy weather situations over Italy*”, are interpreted semantically using a number of problem independent databases and databases that model the domain of the target application. Semantic information extracted from various source documents is aggregated and fused using *fuzzy data fusion algorithms*. With NL the meaning of queries depends heavily on quantifying expressions, as witnessed for example, by the different meaning of *there are few clouds over Italy* vs. *there are lots of clouds over Italy*, which both could be part of queries submitted to our system. Images and all other kinds of data are analysed and indexed either when they are stored in the system or, if there is not enough information in the index available, on-line under a given query. In the meteorological domain, users are typically interested in certain weather conditions in a specified local and temporal region. Weather conditions, however, are *not fully specified by any single document* in the database; for example, satellite images provide the required data on cloudiness, while ground temperature readings can be obtained from temperature maps, this is the case for information fusion across domains. In addition, more than one document may describe the same aspect of a weather situation. Therefore, operators for information combination are required which establish content-related links between the available documents and allow for the fusion of (possibly conflicting) information extracted from several document sources. Fig. 1 shows a typical output of the system. The query submitted in NL was: “Show me pictures of cloud formation over Bavaria in the first week of August”

In this chapter, we present an axiomatic theory of fuzzy quantification that was developed as a sound basis for this system and that has found many more applications beyond this initial use. It is based on the novel concept of a *determiner fuzzification scheme (DFS)*. Unlike existing approaches to fuzzy quantification, DFS is

- a compatible extension of the *theory of generalized quantifiers* (TGQ, Barwise and Cooper (1981));
- a genuine theory of fuzzy multi-place quantification;
- not limited to absolute and proportional quantifiers;
- able to handle both quantitative and non-quantitative (i.e. qualitative) quantifiers;
- not limited to finite universes of discourse;
- based on a rigid axiomatic foundation;
- fully compatible to negation, formation of antonyms, dualisation, and other important constructions on quantifiers.

In the rest of the chapter we explain what fuzzy quantifiers are and why they are of interest to data fusion. We then briefly refer to existing approaches to fuzzy quantification and their drawbacks to go on to describe DFS theory and its advantages over these approaches. After presenting the algorithms needed to compute quantification results based on the DFS models, we finally present examples how fuzzy quantifiers are used in the retrieval system for multimedia weather documents. In particular, we demonstrate the suitability of our approach to fuse a sequence of images under a fusion criterion expressed in natural language.

2 Data fusion and fuzzy quantifiers

There are many statistical approaches to data fusion available, some of them have been applied to real-world tasks with great success (e.g. Brooks and Iyengar; Goodman *et al.*; Hall (1997; 1997;

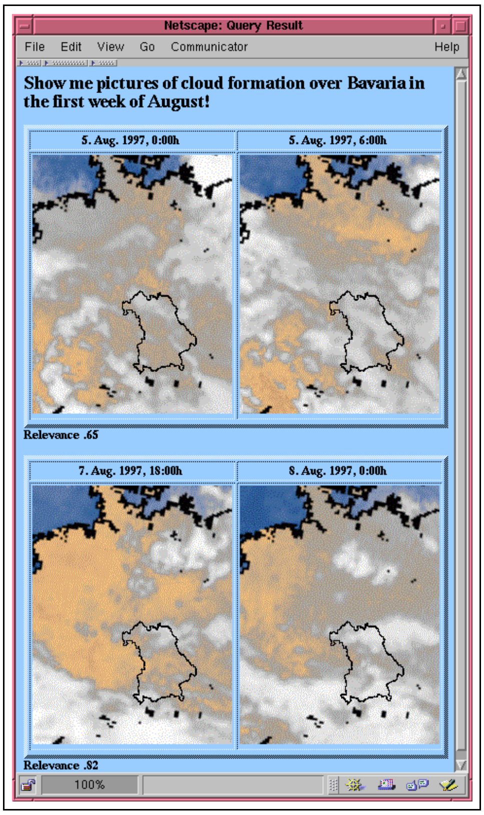

Figure1. Resulting information output of our multimedia database system. For answering the NL query several fuzzy aggregation and fusion operations are necessary.

1992)). They suffer, however, from two main drawbacks: (a) statistical models are normally hard to establish and (b) human knowledge is hard to incorporate into the models and/or the fusion process. Natural language (NL) holds the potential to express this kind of human knowledge. NL would be an ideal candidate for modeling and steering the fusion process because

1. Knowledge about the data sources (sensors etc.) can be expressed in NL statements; e.g. describe conditional reliability of a sensor through **if-then** rules;
2. The fusion criterion can be expressed in NL, e.g. using quantifiers like **all**, **a lot**, **about half** etc.

However, a linguistically adequate and consistent interpretation of NL phrases requires semantic devices to handle a number of difficulties.

– NL concepts (expressed by nouns, verbs, adjectives) often lack clear boundaries.
  Model: use linguistic terms of fuzzy logic

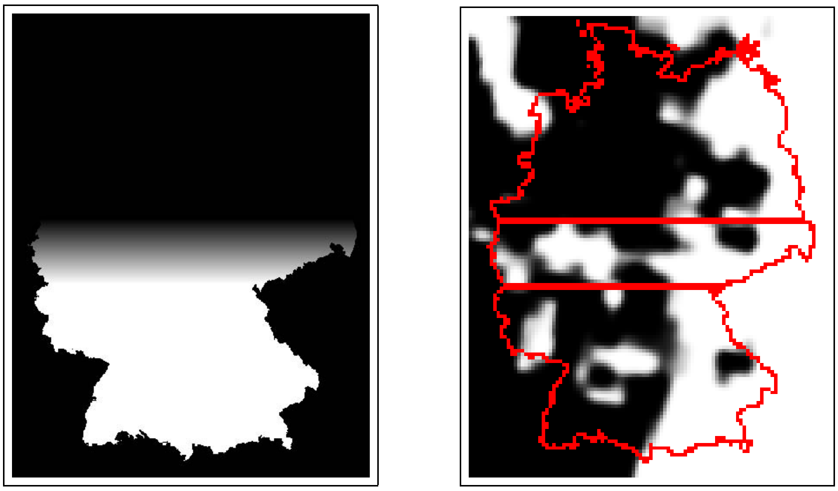

Figure2. Our running example of a base set E, the pixels forming Germany. **Left.** Subset X_1: Southern Germany. **Right.** Subset X_2: Cloudy.

- NL phrases contain modifiers (e.g. **very**, **rather**,...).
  Model: use linguistic hedges of fuzzy logic
- NL quantifiers (e.g. **almost all**, **many**) can express approximate aggregation criteria
  Model: use fuzzy quantifiers

Let us recall some concepts of fuzzy set theory to clarify what is meant by the term "fuzzy quantifier". Suppose that we have some nonempty base set, or domain, E. Fuzzy subsets are useful for modelling concepts like **tall** or **warm** which do not have sharp boundaries. Formally, a fuzzy subset X of the base set E assigns to each element $e \in E$ of the base set a membership grade $\mu_X(e)$ in the unit interval. The set of all fuzzy subsets of E (fuzzy powerset) is denoted by $\widetilde{\mathcal{P}}(E)$. By an n-ary fuzzy quantifier $\widetilde{Q}$ on E we mean a mapping which to each n-tuple of fuzzy subsets $X_1, \ldots, X_n$ of E assigns a gradual quantification result $\widetilde{Q}(X_1, \ldots, X_n) \in [0, 1]$.

Consider the example in Fig. 2. Here, E is a set of pixel coordinates forming the shape of Germany. We have two argument sets: X_1 is a fuzzy subset of E which represents Southern Germany. The region of white pixels fully belongs to Southern Germany, and the grey pixels are those which belong to Southern Germany to some degree. X_2 is a fuzzy subset which represents a cloudiness situation. White pixels are fully cloudy, grey pixels are cloudy to some degree. (The contours of Germany have been added in red in order to facilitate interpretation. The lower part fully belongs to Southern Germany, the upper part belongs to Germany, but not to Southern Germany, and the middle part contains the intermediate cases.) We can then apply a two-place fuzzy quantifier $\widetilde{Q}$ suitable for interpreting **almost all**. Hence, $\widetilde{Q}(X_1, X_2) \in [0, 1]$ is the result of evaluating "Almost all of Southern Germany is cloudy" (Fig. 2, right).

This simple example illustrates some of the key properties of fuzzy quantifiers that make them particularly suited to information fusion: Fuzzy quantifiers are *powerful fusion operators* which can model the *weighted* aggregation of sets of gradual criteria (compare to simple Boolean

conjunction	k_1 **and** ... **and** k_m corresponds to: **all** criteria $k_1, \ldots, k_m$ are satisfied
disjunction	k_1 **or** ... **or** k_m corresponds to: **at least one** criterion $k_1, \ldots, k_m$ is satisfied
fuzzy quantifiers:	$\widetilde{Q}(W, G)$ corresponds to: $\widetilde{Q}$ **important** criteria are **satisfied** $\widetilde{Q} \in \{$**almost all**, **many**, **about ten**, ...$\}$: fuzzy quantifier W: degree of importance G: degree of validity $\widetilde{Q} =$ **all** models weighted conjunction, $\widetilde{Q} =$ **at least one** yields weighted disjunction.

Table1. Fuzzy quantifiers and weighted aggregation

operators like **and**, **or**, **not**). Table 1 illustrates some correspondences between fuzzy quantifiers and the NL meaning of connectives: Assuming that we have a set E of criteria, a fuzzy subset W of important criteria and a fuzzy subset G of the grades to which the criteria are satisfied, then every two-place quantifier $\widetilde{Q}$ defines a weighted aggregation criterion "$\widetilde{Q}$ important criteria are satisfied", which can be evaluated by computing $\widetilde{Q}(W, G)$, where $\widetilde{Q}$ is, for example, **almost all**, **many** etc. In Glöckner and Knoll (2000), it has been shown that this approach provides a natural account of weighted aggregation that avoids the misconceptions inherent to the well-known Waller-Kraft wish list of weighted Boolean information retrieval, introduced in Waller and Kraft (1979). Moreover, fuzzy quantifiers are intuitively familiar to users of information systems, and can be applied for technical fusion purposes in the same way as in everyday language (unlike, say, the fuzzy integral, elaborate methods from statistics, Dempster-Shafer theory, etc.). Bordogna and Pasi (1997) describe a promising application of fuzzy quantifiers to information retrieval, which are used to fuse matches of the search terms across the available search fields (i.e. document sections like title, abstract, conclusion). The fusion criterion is controlled linguistically by selecting an appropriate quantifier; it can be further refined by assigning weights of importance to the individual sections. Apparently, the success of such applications of fuzzy quantifiers to information fusion is highly dependent on the model of quantification used, which must be linguistically plausible in order to ensure that the computed results convey the intended semantics.

3 Traditional approaches to fuzzy quantification

Now let us turn to the interesting question of which fuzzy quantifier $\widetilde{Q}$ corresponds to a given natural language quantifier like **almost all**. To solve this problem, Zadeh (1979) proposed a simple representation of fuzzy quantifiers. Absolute quantifiers like "about ten" are modelled as fuzzy subsets of the non-negative reals with membership functions $\mu_Q : \mathbb{R}_0^+ \longrightarrow [0, 1]$, while proportional quantifiers like "most" or "more than 30 percent" are represented as fuzzy subsets of the unit interval with membership functions $\mu_Q : [0, 1] \longrightarrow [0, 1]$. A possible definition of the proportional quantifier "almost all" is shown in Fig. 3.

In order to make these fuzzy numbers applicable to fuzzy sets for the purpose of quantification, a mechanism (which we denote by $\mathcal{Z}$) is needed which maps μ_Q to a fuzzy quantifier

$\mathcal{Z}^{(1)}(\mu_Q) : \widetilde{\mathcal{P}}(E) \longrightarrow [0,1]$, in order to model the unrestricted use of the quantifier, relative to E, or $\mathcal{Z}^{(2)}(\mu_Q) : \widetilde{\mathcal{P}}(E)^2 \longrightarrow [0,1]$, in order to model its restricted use, relative to the first argument. (Additional subscripts “abs” and ‘prp” will be used in order to distinguish absolute quantifiers from proportional quantifiers). An example of the unrestricted use of a quantifier is “most (elements of the domain) are tall”. In this case, the interpretation solely depends on the fuzzy subset **tall** $\in \widetilde{\mathcal{P}}(E)$ of tall people, and is accomplished by computing $\mathcal{Z}^{(1)}_{\mathrm{prp}}(\mu_{\textbf{most}})(\textbf{tall})$. By contrast, “most young people are tall” illustrates the restricted use of a quantifier. In this case, the interpretation depends upon two fuzzy argument sets, viz the fuzzy set **young** $\in \widetilde{\mathcal{P}}(E)$ of young people and the fuzzy set **tall** $\in \widetilde{\mathcal{P}}(E)$ of tall people. The interpretation is hence accomplished by evaluating $\mathcal{Z}^{(2)}_{\mathrm{prp}}(\mu_{\textbf{most}})(\textbf{young}, \textbf{tall})$.

Various approaches have been proposed to transform the simplified representation μ_Q into a fuzzy quantifier, viz the Σ-count approach of Zadeh (1983), the OWA approach of Yager (1988) which is based on ordered weighted averaging (OWA) operators, the FG-count approach of Zadeh (1983) and the FE-count approach of Ralescu (1986). Each of these approaches provides its unique definition of the above mechanism $\mathcal{Z}$.

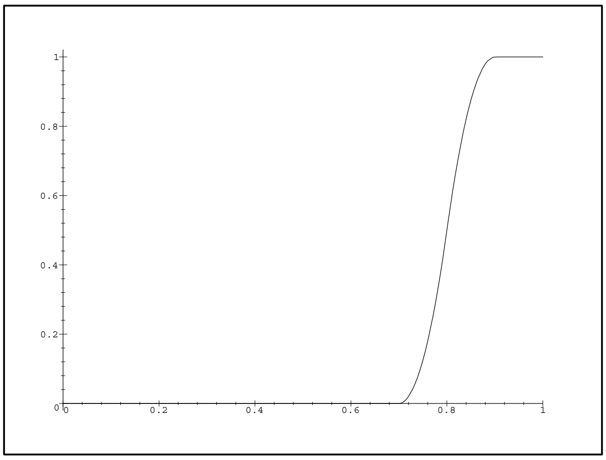

Figure3. A possible definition of the proportional quantifier **almost all**.

Zadeh's Σ-count approach is based on the computation of Σ-counts of a quantifier's arguments. These are a scalar measure of the cardinality of fuzzy sets, defined as the sum of all membership grades, viz $\Sigma\text{-count}(X) = \sum_{e \in E} \mu_X(e)$. In our notation, Zadeh's definitions of the Σ-count approach give rise to the following choice of the mechanism $\mathcal{Z}$.

Interpretation of absolute quantifiers, $\mu_Q : \mathbb{R}_0^+ \longrightarrow [0,1]$.

One-place quantification: $\mathrm{SC}^{(1)}_{\mathrm{abs}}(\mu_Q)(X) = \mu_Q\,(\Sigma\text{-count}(X))$

Two-place quantification: $\mathrm{SC}^{(2)}_{\mathrm{abs}}(\mu_Q)(X_1, X_2) = \mathrm{SC}^{(1)}_{\mathrm{abs}}(\mu_Q)(X_1 \cap X_2)$

Interpretation of proportional quantifiers, $\mu_Q : [0,1] \longrightarrow [0,1]$.

One-place quantification: $\mathrm{SC}^{(1)}_{\mathrm{prp}}(\mu_Q)(X) = \mathrm{SC}^{(2)}_{\mathrm{prp}}(\mu_Q)(E, X) = \mu_Q\,(\Sigma\text{-count}(X)/|E|)$

Two-place quantification: $\mathrm{SC}^{(2)}_{\mathrm{prp}}(\mu_Q)(X_1, X_2) = \mu_Q\left(\dfrac{\Sigma\text{-count}(X_1 \cap X_2)}{\Sigma\text{-count}(X_1)}\right)$

Hence an unary absolute quantifier is computed by substituting X into the Σ-count. Let $X =$ **young** be the fuzzy set of "young" persons and $s = \Sigma$-count (**young**) their number as measured by Σ-count and let μ_Q be the membership function of an absolute quantifier (e.g. **about 10'**) then "about ten (elements of the domain) are young" is computed as $\mu_Q(s)$. Two-place absolute quantifiers, e.g. "about ten of the young are tall" are transformed into an unary quantifier plus conjunction "about ten (elements of the domain) are young AND tall" and computed as $\mu_Q(s)$, where $s = \Sigma$-count (**young** $\cap$ **tall**). For proportional two-place quantifiers the relative share of the X_1 that *are* X_2 is computed, i.e. $s = \Sigma$-count (**young** $\cap$ **tall**)/ Σ-count (**young**).

One drawback of the Σ-count approach is the following: If the natural language quantifier to be modelled is two-valued, then the Σ-count approach produces two-valued fuzzy quantifiers, i.e. operators the results of which are always 1 (fully true) or 0 (fully false). This is undesirable because such operators are discontinuous, i.e. very sensitive to slight changes in the arguments of the quantifier. The example in Fig. 4 illustrates this extreme sensitivity of the Σ-count approach: Although the cloudiness situations depicted in the center and the right subfigure are very similar, the results obtained from the Σ-count approach are totally different (0 vs. 1). It has been shown in Glöckner (1999) that it is not permissible to replace μ_Q with a smoothened choice of quantifier in order to avoid these effects, because the intended semantics of the quantifier would be lost.

A further disadvantage of the Σ-count approach is that a large number of "small" membership grades can accumulate in an undesirable way, see Fig. 5. In the center result image there is about ten percent cloud coverage of Southern Germany, so a result of 1 is as expected. In the right image, however, *all* of Southern Germany is almost not cloudy (i.e. cloudy to the low degree of 10%) which does not mean that ten percent of Southern Germany are cloudy. If one claims that 10 percent of Southern Germany are cloudy in the right image, one should be able to tell *which* 10% of Southern Germany is cloudy. But there is no reasonable choice because all pixels which belong to Southern Germany are almost not cloudy.

Another prominent approach to fuzzy quantification is Yager's approach based on ordered weighted averaging (OWA) operators, see Yager (1988) and Yager (1991). In this case, only *proportional quantifiers* like "about one half" are considered. In addition, the quantifier is assumed to be 'regular nondecreasing', i.e. $\mu_Q : [0,1] \longrightarrow [0,1]$ satisfies $\mu_Q(0) = 0$, $\mu_Q(1) = 1$ and $\mu_Q(x) \leq \mu_Q(y)$ whenever $x \leq y$. The definitions proposed by Yager correspond to the following choice of quantification mechanism $\mathcal{Z}$. Let $m = |E|$ be the number of elements in the basic domain. Then $\mu_Q(j/m) - \mu_Q((j-1)/m)$ is the quantifier's increment if the number of elements $j - 1$ is increased to j, or proportionally from $(j-1)/m$ to j/m. By $\mu_{[j]}(X)$ we denote the j-th greatest membership grade of X (including duplicates), i.e. the j-th element in the

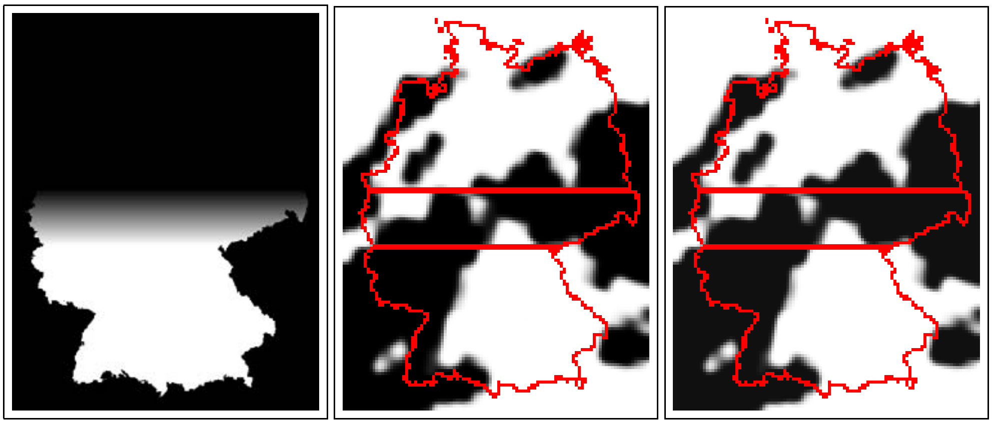

Figure4. Results of computing "**At least 60 percent of Southern Germany are cloudy"** with the Σ-count approach. Left: Southern Germany. Result for center image: 0, result for right image: 1.

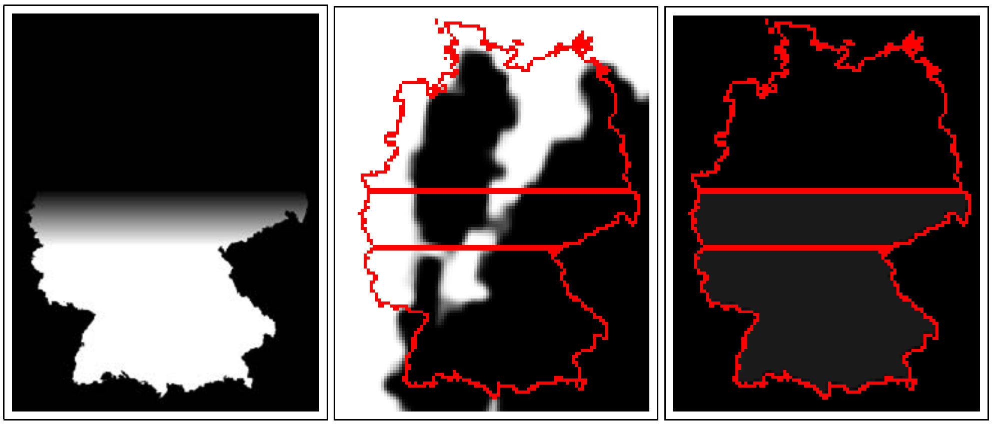

Figure5. The Σ-count approach accumulates "small" membership grades in an undesirable way. Results of **"About 10 percent of Southern Germany are cloudy"**. Left: Southern Germany. Result for center image: 1 (which is plausible). Result for right image: 1 (totally inadequate).

ordered sequence of membership values. Intuitively, $\mu_{[j]}(X)$ expresses the grade at which X has at least j elements. Originally, for one-place quantification, Yager proposed the following rule for computation of the quantifier:

$$\mathrm{OWA}^{(1)}_{\mathrm{prp}}(\mu_Q)(X) = \sum_{j=1}^{m} (\mu_Q(j/m) - \mu_Q((j-1)/m)) \cdot \mu_{[j]}(X)$$

In other words: the quantifier value increment when increasing cardinality from $j-1$ to j is weighted by the grade at which X has at least j elements. This is repeated for all j and the results are summed up. This results in good adequacy with 'regular nondecreasing' quantifiers and one-place quantification, but it is relatively limited in applications.

The extension to two-place quantification proposed in Yager (1991) transforms the fuzzy arguments X_1, X_2 (depending upon the quantifier) into a single fuzzy set Z to which the unary quantification is applied:

$$\mathrm{OWA}^{(2)}_{\mathrm{prp}}(\mu_Q)(X_1, X_2) = \mathrm{OWA}^{(1)}_{\mathrm{prp}}(\mu_Q)(Z)$$

where Z is defined by

$$\mu_Z(e) = \max(\mu_{X_1}(e), 1 - \mathrm{orness}(\mu_Q)) \cdot \mu_{X_2}(e)^{\max(\mu_{X_1}(e), \mathrm{orness}(\mu_Q))}$$

and

$$\mathrm{orness}(\mu_Q) = \frac{1}{m-1} \sum_{j=1}^{m-1} \mu_Q(j/m)\,.$$

The "degree of orness" is a formal coefficient, which affects the computation of Z. For "all" it is 0, for "exists" it is 1, and for the other quantifiers its value is in between.

This formula yields the correct results for "all" and "at least one". However, it has been shown by Glöckner and Knoll (1999a) that it fails in the case of all other proportional quantifiers. Note that it is not the form of OWA that was proposed for unary regular nondecreasing quantifiers, it is the generalisation to two-place quantifiers that induces problems as illustrated in the following example: As shown in Fig. 6, the task is to determine the degree to which "At least 60 percent of Southern Germany are cloudy". This is clearly the case in the center image, but not in the right image. The OWA approach, however, ranks the right image higher than the center image, which is clearly implausible. As testified by Fig. 6 results of the OWA approach show undesirable dependency on cloudiness grades in regions III and IV, which do not belong to Southern Germany at all. This is because the two-place OWA formula cannot model any *conservative* quantifiers beyond **all** and **exists**, see Glöckner (1999).

The proposed formula leads to completely implausible effects for other quantifications, too. Examples are listed in Table 3 (referring to Fig. 6). The quantifier "less than 60 percent" fails to be regular nondecreasing. Hence it cannot be modelled by OWA directly. There are two natural language paraphrases which reduce "less than 60 percent" to a regular nondecreasing quantifier. We can use the negation "at least 60 percent" and evaluate "less than 60 percent of the X_1's are X_2's" by negating the result of "at least 60 percent of the X_1's are X_2's". Alternatively, we can use the antonym "more than 40 percent of the X_1's are X_2's" and evaluate "less than 60 percent of the X_1's are X_2's" by computing "more than 40 percent of the X_1's are (NOT X_2)'s". Results should coincide in both cases because they stem from equivalent NL paraphrases, but they differ when applying the two-place OWA formula, which is *not compatible to dualisation*.

Another approach to fuzzy quantification was proposed by Zadeh (1983). It is based on the computation of the FG-count, which is a fuzzy measure of the cardinality of a fuzzy set. The 'basic' FG-count approach is restricted to absolute quantifiers and proportional one-place quantifiers. The definition, only applicable to nondecreasing μ_Q, is as follows:

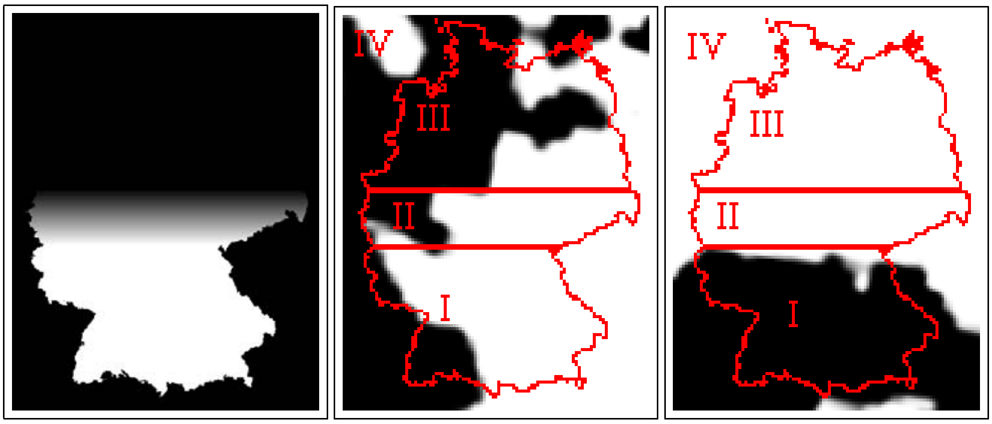

Figure6. Application of the OWA-approach. Results of **"At least 60 percent of Southern Germany are cloudy"**. Left: Southern Germany. Result for center image: OWA: 0.1, desirable outcome: 1. Result for right image: OWA: 0.6, desirable: 0.

Interpretation of absolute quantifiers, $\mu_Q : \mathbb{R}_0^+ \longrightarrow [0,1]$**.**

One-place quantification: $\mathrm{FG}^{(1)}_{\mathrm{abs}}(\mu_Q)(X) = \max\{\min(\mu_Q(j), \mu_{[j]}(X)) : j = 0, \ldots, |E|\}$

Two-place quantification: $\mathrm{FG}^{(2)}_{\mathrm{abs}}(\mu_Q)(X_1, X_2) = \mathrm{FG}^{(1)}_{\mathrm{abs}}(X_1 \cap X_2)$

Interpretation of proportional quantifiers, $\mu_Q : [0,1] \longrightarrow [0,1]$**.**

One-place quantification: $\mathrm{FG}^{(1)}_{\mathrm{prp}}(\mu_Q)(X) = \max\{\min(\mu_Q(j/m), \mu_{[j]}(X)) : j = 0, \ldots, |E|\}$

FG-Count(X) is a fuzzy subset of the natural numbers where $\mu_{\text{FG-Count}(X)}(j) = \mu_{[j]}(X)$ for all $j \in \mathbb{N}$. (Here and in the following it is convenient to stipulate that $\mu_{[0]}(X) = 1$ and $\mu_{[j]}(X) = 0$ for $j > |E|$.) Hence FG-Count(X) defines for a given j, to which degree X has at least j elements. For a given number j of elements in the set X the relation $\min(\mu_Q(j), \mu_{[j]}(X))$ computes the degree to which the quantifier applies to j AND X has at least j elements. Then we maximise the results for every choice of j. This basic approach is well-behaved but rather limited, because it is only defined for nondecreasing quantifers, and cannot handle the important case of two-place quantification with proportional quantifiers.

Quantifier	center image	right image
"not at least 60 percent"	0.9	0.4
"more than 40 percent not"	0.4	0
desired result	0	1

Table2. Results of other OWA-quantifiers when applied to the images in Fig. 6

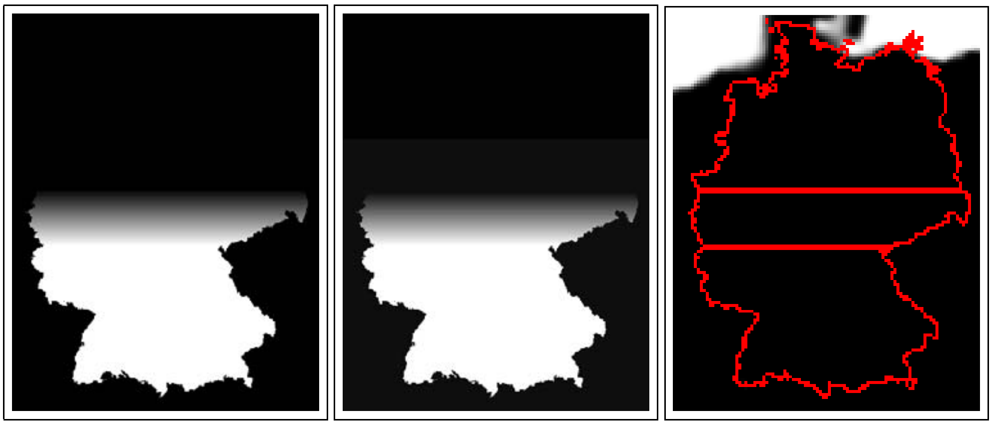

Figure7. Application of the FG-approach. Results of **"At least 5 percent of Southern Germany are cloudy"**. Left: Southern Germany (a). Center image: Southern Germany (b). Right image: Cloudy. Result for Southern Germany (a): 0.55, Result for Southern Germany (b): 0.95. The desired result: 0.

Yager (1991, p. 72) has suggested a formula which extends the FG-count approach to the case of proportional two-place quantification.

$$\mathrm{FG}^{(2)}_{\mathrm{prp}}(\mu_Q)(X_1, X_2) = \max\left\{\min\left(\mu_Q\left(\frac{\sum_{v\in S}\mu_{X_1}(v)}{\sum_{v\in E}\mu_{X_1}(v)}\right), H_S\right) : S \in \mathcal{P}(E)\right\}$$
$$H_S = \min\{\max(1 - \mu_{X_1}(v), \mu_{X_2}(v)) : v \in S\}\,.$$

To see that this formula is related to the FG-count approach, suffice to observe that

$$\mathrm{FG}^{(1)}_{\mathrm{prp}}(\mu_Q)(X) = \mathrm{FG}^{(2)}_{\mathrm{prp}}(\mu_Q)(E, X)$$

for all μ_Q and $X \in \widetilde{\mathcal{P}}(E)$, if we use the above formula to evaluate $\mathrm{FG}^{(2)}_{\mathrm{prp}}$.

Here, as in the case of OWA, the generalisation to two-place quantification is not plausible. If we define a condition that "at least five percent of Southern Germany are cloudy", this can be illustrated using our standard images (see Fig. 7). There are no clouds in Southern Germany (a) at all but the FG-count result is 0.55. This implausible behaviour of the formula for two-place proportional quantifiers is caused by its *lack of preserving local monotonicity properties* of quantifiers, see Glöckner (1999). A jump of result from 0.55 to 0.95 when moving to Southern Germany (b), a slightly modified more noisy image, illustrates that the resulting *FG-count quantifiers can be discontinuous*, i.e. very sensitive to noise and hence not suited for practical applications. A similar behaviour as with OWA can be provoked when quantifications like "Less than 30 percent of Southern Germany are cloudy", which fail to be to be nondecreasing, are replaced with NL paraphrases based on nondecreasing quantifiers: (a) "It is not the case that at least 30 percent of Southern Germany are cloudy", i.e. the result for negated quantifier is computed and this result is negated; and (b) "More than 70 percent of Southern Germany are not cloudy", which makes use of the antonym. These NL paraphrases are equivalent and should be interchangeable. The two-place FG-formula, however, generates different results in both cases.

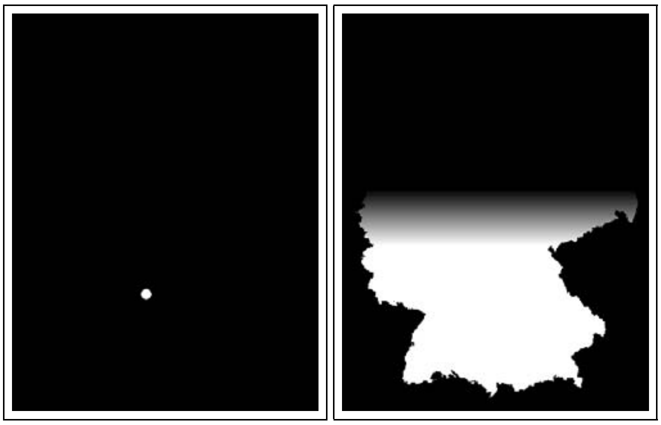

Figure8. Application of the FE-approach. Results of **"The image region X is nonempty"**. Left: desired: 1, FE: 1. Right: desired: 1, FE: 0.5.

We finally refer to the FE-count approach, proposed by Ralescu (1986) based on the computation of FE-counts, another fuzzy measure of the cardinality of fuzzy sets proposed by Zadeh (1983). To lift the restrictions to nondecreasing quantifiers with the FG-count approach, the FG-count is replaced with the FE-count. The basic idea is that $\mu_{[j]}(X)$, i.e. the j-th greatest membership grade of the fuzzy set X, only specifies the membership degree at which X contains *at least* j elements and hence does not fully capture the notion of cardinality. It is therefore replaced with $\min(\mu_{[j]}(X), 1 - \mu_{[j+1]}(X))$, which computes the degree at which X has at least j, but not $j+1$ elements, i.e. the grade at which X has exactly j elements. For a given j this grade is combined conjunctively with the evaluation of the quantifier $\mu_Q(j)$ by "min". The individual results are maximised for all choices of j in order to get the final value. Here is the definition of one-place quantification for absolute quantifiers $\mu_Q : \mathbb{N} \longrightarrow [0, 1]$, based on the FE-count:

$$\mathrm{FE}^{(1)}_{\mathrm{abs}}(\mu_Q)(X) = \max\{\min\{\mu_{[j]}(X), 1 - \mu_{[j+1]}(X), \mu_Q(j)\} : j = 0, \ldots, |E|\}\,.$$

In the example, the image regions in the left and right image of Fig. 8 are clearly nonempty. The left image contains a crisp nonempty region. The fuzzy image region in the right image is certainly nonempty to the degree of 1, too, because it contains the crisp nonempty image region in the left image. The FE-count approach, however, only rates the left image correctly as (fully) nonempty. With the larger region in the right image, the result drops to 0.5, which is clearly inacceptable. The example hence reveals a severe misconception of the FE-count approach, which fails to preserve monotonicity properties of quantifiers.

What are the reasons of the failure of the quantifiers presented above, why are existing approaches to fuzzy quantification poor models of NL quantification and hence unsuitable for our goal of capitalising of NL for data fusion? The answer is manifold, but the most important reasons result from fuzzy set theory having ignored what linguists know about NL quantifiers:

- There is a well-established linguistic theory of NL quantification: the Theory of Generalized Quantifiers (TGQ), see e.g. Barwise and Cooper (1981) and van Benthem (1983);

– TGQ has developed concepts which describe important aspects of the meaning of NL quantifiers, *e.g. negation, antonyms, duality, monotonicity properties, symmetry, having extension, conservativity, adjectival restriction...*
– the counter-intuitive behaviour of existing approaches can be explained in terms of an incompatibility to concepts of TGQ.
– compared to TGQ, existing approaches can model only a very limited class of NL quantifiers, i.e. fuzzy numbers with membership functions $\mu_Q : \mathbb{R}_0^+ \longrightarrow [0, 1]$ and $\mu_Q : [0, 1] \longrightarrow [0, 1]$ are too weak a representation.

4 A new approach to fuzzy quantification: DFS

As argued in the previous section, the traditional approaches fail, mostly in the case of n-place quantification (with $n > 1$), even in the simplest situations. TGQ, which is an elaborate linguistic theory of NL quantification, by contrast, defines n-ary quantifiers such as "more A than B are C", "most A's and B's are C's or D's", including non-quantitative quantifiers like "John" or "almost all good X's are Y's". According to Keenan and Stavi (1986) there are more than 30 different types, of which absolute and proportional quantifiers are only 2 basic examples. Nonetheless, TGQ was not developed with fuzzy sets in mind.

The incompatibility of the traditional approaches with the concepts developed in TGQ is the main reason for their poor performance: lack of compatibility with conservative quantifiers and dualisation in the case of OWA; incompatibility with local monotonicity properties, dualisation, and possible discontinuities in the case of the FG-count approach; failure to preserve monotonicity properties with the FE-count approach. The Σ-count approach is more difficult to assess, but it does have problems with accumulation and it may result in discontinuous operators.

4.1 An overview of DFS

What can a better approach to fuzzy quantification look like? We identify a number of cornerstones of a principled theory:

1. *Improved Representation through n-ary quantifiers and semi-fuzzy quantifiers*
   We recall that an n-ary fuzzy quantifier $\widetilde{Q}$ on a base set E assigns to each choice of fuzzy subsets $X_1, \ldots, X_n$ of E a gradual result $\widetilde{Q}(X_1, \ldots, X_n) \in [0, 1]$. Fuzzy quantifiers constitute an expressive class of operators because they introduce a second order construct for fuzzy sets. However, they are often hard to define because the familiar concept of cardinality of crisp sets is not applicable to the fuzzy sets that form the arguments of a fuzzy quantifier. It is therefore necessary to introduce a simplified representation, which must be still powerful enough to embed all quantifiers in the sense of TGQ. This representation is provided by *semi-fuzzy quantifiers*. An n-ary *semi-fuzzy quantifier* on a base set E is a mapping $Q : \mathcal{P}(E)^n \longrightarrow [0, 1]$ which to each choice of *crisp* subsets $Y_1, \ldots, Y_n$ of E assigns a gradual result $Q(Y_1, \ldots, Y_n) \in [0, 1]$. Because semi-fuzzy quantifiers must be defined for crisp input only, they are much easier to define than fuzzy quantifiers. In particular, the usual crisp cardinality is applicable to their arguments and can hence be used to provide an interpretation for semi-fuzzy quantifiers. (Examples of semi-fuzzy quantifiers will be given in sections 4.5 and 5).

2. A *quantifier fuzzification mechanism* (QFM) $\mathcal{F}$ assigns to each semi-fuzzy quantifier Q a fuzzy quantifier $\mathcal{F}(Q)$ of the same arity and on the same base set. These are applicable both to crisp and fuzzy arguments. QFMs are useful because the concepts of TGQ can be easily adapted to the case of semi-fuzzy quantifiers and fuzzy quantifiers. We can then require that a certain property of a quantifier be preserved when applying the QFM, and that $\mathcal{F}$ be compatible with certain constructions on (semi-)fuzzy quantifiers. This can be likened to the well-known mathematical concept of a homomorphism (structure-preserving mapping).
3. We require *compatibility* with concepts of TGQ by stating *axioms for "admissible" or "reasonable"* choices of $\mathcal{F}$. An adequate QFM should preserve all properties of linguistic relevance. We enforce this by stating a set of axioms for 'admissable' or 'reasonable' choices of QFMs, the DFS axioms.
4. We should find *models of the axioms*, i.e. "reasonable" choices of $\mathcal{F}$, and characterise interesting classes of such models in terms of distinguished properties;
5. *Efficient algorithms* must be developed for implementing the resulting operators.

Following from these requirements, we have developed an axiomatic theory of fuzzy quantification known as "DFS theory". We start the description of our theory by listing the 6 axioms that are required:

1. **Correct Generalisation.** We require that

$$\mathcal{F}(Q)(X_1, \ \dots, \ X_n) = Q(X_1, \ \dots, \ X_n)$$

whenever $X_1, \dots, X_n \in \mathcal{P}(E)$ are crisp (combined with the other axioms, this condition can be restricted to $n \leq 1$).
Rationale: a semi-fuzzy quantifier Q is defined only for crisp arguments, while $\mathcal{F}(Q)$ is defined for arbitrary fuzzy arguments. If all arguments are crisp, Q and $\mathcal{F}(Q)$ must match.
2. **Membership Assessment.** The two-valued quantifier defined by $\pi_e(Y)$ =1 if $e \in Y$ and $\pi_e(Y) = 0$ otherwise for $Y \in \mathcal{P}(E)$, has the obvious fuzzy counterpart $\widetilde{\pi}_e(X) = \mu_X(e)$, $X \in \widetilde{\mathcal{P}}(E)$. We require that $\mathcal{F}(\pi_e) = \widetilde{\pi}_e$.
Rationale: Membership assessment (crisp or fuzzy) can be modelled through quantifiers. For an element e of the base set, we can define a two-valued quantifier π_e which checks if e is present in its argument. Similarly, we can define a fuzzy quantifier $\widetilde{\pi}_e$ which returns the degree to which e is contained in its argument. It is natural to require that the crisp quantifier π_e be mapped to $\widetilde{\pi}_e$, which plays the same role in the fuzzy case.
3. **Dualisation.** We require that $\mathcal{F}$ preserves dualisation of quantifiers, i.e.

$$\mathcal{F}(Q')(X_1, \ \dots, \ X_n) = \widetilde{\neg}\, \mathcal{F}(Q)(X_1, \dots, X_{n-1}, \widetilde{\neg}\, X_n)$$

for all fuzzy arguments $X_1, \dots, X_n$ whenever $Q'(Y_1, \ \dots, \ Y_n) = \widetilde{\neg}\, Q(Y_1, \dots, Y_{n-1}, \neg Y_n)$ for all crisp arguments $Y_1, \dots, Y_n$.
Rationale: Obviously, a phrase like "all X's are Y's" should have the same result as "it is not the case that some X's are not Y's".
4. **Union.** We require that $\mathcal{F}$ preserves unions of arguments, i.e.

$$\mathcal{F}(Q')(X_1, \dots, X_{n+1}) = \mathcal{F}(Q)(X_1, \dots, X_{n-1}, X_n \,\widetilde{\cup}\, X_{n+1})$$

whenever $Q'(Y_1, \ldots, Y_{n+1}) = Q(Y_1, \ldots, Y_{n-1}, Y_n \cup Y_{n+1})$.
Rationale: It should not matter whether "many X's are Y's or Z's" is computed by evaluating $\mathcal{F}(\textbf{many})(X, Y \widetilde{\cup} Z)$ or by computing $\mathcal{F}(Q)(X, Y, Z)$ with $Q(X, Y, Z) = \textbf{many}(X, Y \cup Z)$.

5. **Monotonicity in Arguments.** We require that $\mathcal{F}$ preserve monotonicity in arguments, i.e. if Q is nondecreasing/nonincreasing in the i-th argument, then $\mathcal{F}(Q)$ has the same property. When combined with the other axioms, the condition can be restricted to the case that Q is nonincreasing in its n-th argument).
Rationale: There must be a systematically different interpretation of statements like "all men are tall" and "all young men are tall" where the former statement expresses the stricter condition.
6. **Functional Application.** We require that

$$\mathcal{F}(Q')(X_1, \ \ldots, \ X_n) = \mathcal{F}(Q)(f'_1(X_1), \ldots, f'_n(X_n))$$

whenever Q' is defined by $Q'(Y_1, \ldots, Y_n) = Q(f_1(Y_1), \ldots, f_n(Y_n))$, where $f'_1, \ldots, f'_n$ are obtained from the induced extension principle of $\mathcal{F}$, see Glöckner (2000a).
Rationale: This abstract axiom ensures that $\mathcal{F}$ behave consistently over different domains E.

A QFM $\mathcal{F}$ which satisfies these axioms is called a determiner fuzzification scheme, or DFS for short ("determiner" is a synonym from TGQ for "quantifier"). If $\mathcal{F}$ induces the standard negation $\neg x = 1 - x$ and the standard extension principle of Zadeh (1975), then it is called a *standard DFS*. These DFSes constitute the natural class of standard models of fuzzy quantification.
A large number of properties of linguistic or logical relevance are entailed by the above axioms: If $\mathcal{F}$ is a DFS, then

- $\mathcal{F}$ *induces a reasonable set of fuzzy propositional connectives*, i.e. $\widetilde{\neg}$ is a strong negation $\widetilde{\wedge}$ is a t-norm, $\widetilde{\vee}$ is an s-norm etc. These connectives are obtained from a canonical construction, see Glöckner (2000a);
- $\mathcal{F}$ *is compatible with the negation of quantifiers*, i.e.

$$\mathcal{F}(Q')(X_1, \ \ldots, \ X_n) = \widetilde{\neg}\, \mathcal{F}(Q)(X_1, \ \ldots, \ X_n)$$

  whenever $Q'(Y_1, \ \ldots, \ Y_n) = \widetilde{\neg}\, Q(Y_1, \ \ldots, \ Y_n)$. For example, the meanings of "at least one tall men is lucky" and "it is not the case that no tall man is lucky" coincide in every DFS;
- $\mathcal{F}$ *is compatible with the formation of antonyms*, i.e.

$$\mathcal{F}(Q')(X_1, \ \ldots, \ X_n) = \mathcal{F}(Q)(X_1, \ldots, X_{n-1}, \widetilde{\neg}\, X_n)$$

  whenever $Q'(Y_1, \ \ldots, \ Y_n) = Q(Y_1, \ldots, Y_{n-1}, \neg Y_n)$. For example, the meanings of "every tall men is bald" and "no tall men is not bald" coincide in every DFS;
- $\mathcal{F}$ *is compatible with intersections*, i.e.

$$\mathcal{F}(Q')(X_1, \ldots, X_{n+1}) = \mathcal{F}(Q)(X_1, \ldots, X_{n-1}, X_n \,\widetilde{\cap}\, X_{n+1})$$

  whenever $Q'(Y_1, \ldots, Y_{n+1}) = Q(Y_1, \ldots, Y_{n-1}, Y_n \cap Y_{n+1})$. For example, the meanings of "at least two X's are Y's" and "the set of X's that are Y's contains at least two elements" coincide in every DFS;

– $\mathcal{F}$ *is compatible with argument permutations*, i.e.

$$\mathcal{F}(Q')(X_1, \ldots, X_n) = \mathcal{F}(Q)(X_{\beta(1)}, \ldots, X_{\beta(n)})$$

whenever $Q'(Y_1, \ldots, Y_n) = Q(Y_{\beta(1)}, \ldots, Y_{\beta(n)})$, where β is a permutation of $\{1, \ldots, n\}$. In particular, symmetry properties of a quantifier are preserved by applying $\mathcal{F}$. Hence the meaning of "about 50 X's are Y's" and "about 50 Y's are X's" coincide in every DFS.

– $\mathcal{F}$ *is compatible with argument insertion*, i.e.

$$\mathcal{F}(Q')(X_1, \ldots, X_n) = \mathcal{F}(Q)(X_1, \ldots, X_n, A)$$

whenever $Q'(Y_1, \ldots, Y_n) = Q(Y_1, \ldots, Y_n, A)$ for a fixed crisp argument $A \in \mathcal{P}(E)$. For example, the meanings of "many (married X)'s are Y's" and "(many married) X's are Y's" coincide in every DFS.

There is a number of further important properties that conform to previous theoretical work on fuzzy quantification, and reflect some additional concepts of TGQ: If $\mathcal{F}$ is a DFS, then

– $\mathcal{F}(\forall)$ *is a T-quantifier and $\mathcal{F}(\exists)$ is an S-quantifier in the sense of Thiele (1994).* This means that the universal quantifier $\forall$ and the existential quantifier $\exists$ are interpreted plausibly in every DFS.

– $\mathcal{F}(Q)$ *is quantitative if and only if Q is quantitative,*where we assume the usual definition of quantitativity in terms of automorphism-invariance. Hence Q is said to be *quantitative* if for every permutation $\beta : E \longrightarrow E$ of the base set: $Q(\beta(Y_1), \ldots, \beta(Y_n)) = Q(Y_1, \ldots, Y_n)$. The corresponding definition for fuzzy quantifiers is straightforward. Because the result of Q is invariant under permutations of the elements of E, Q cannot rely on any special properties of the elements and only depends on quantitative properties of the arguments. A DFS is guaranteed to map quantitative quantifiers like **almost all** or **a few** to quantitative fuzzy quantifiers; and it is guaranteed to map non-quantitative quantifiers like **John** or **most married** to non-quantitative fuzzy quantifiers.

– $\mathcal{F}$ *is contextual*, i.e. if Q, Q' are defined on a base set E and $X_1, \ldots, X_n \in \widetilde{\mathcal{P}}(E)$ is a choice of fuzzy arguments with $Q(Y_1, \ldots, Y_n) = Q'(Y_1, \ldots, Y_n)$ for all choices of crisp Y_i with $\mathrm{core}(X_i) \subseteq Y_i \subseteq \mathrm{support}(X_i)$, $i = 1, \ldots, n$, then $\mathcal{F}(Q)(X_1, \ldots, X_n) = \mathcal{F}(Q')(X_1, \ldots, X_n)$. Here $\mathrm{core}(X_i)$ denotes the core of the fuzzy subset, i.e. the set of elements with a unity grade of membership. $\mathrm{support}(X_i)$ denotes the support of X_i, i.e. the set of elements with non-zero membership. Apparently, a fuzzy subset is ambiguous only with respect to those elements which are contained in the support, but not in the core: if an element is not contained in the support, then it is fully outside X, while an element contained in the core is fully contained in X. Therefore the quantification result obtained for $\mathcal{F}(Q)(X_1, \ldots, X_n)$ should only depend on the behaviour of Q inside this ambiguity range. In other words, if Q and Q' coincide within the context (ambiguity range) of the fuzzy arguments, then we should obtain the same results. This important property of plausible models, called "contextuality", is possessed by every DFS, see Glöckner (2000a, Theorem 75).

– $\mathcal{F}$ *preserves extension*, i.e. if $E \subseteq E'$ are given base sets, Q is defined on E, Q' is defined on E' and $Q'(Y_1, \ldots, Y_n) = Q(Y_1, \ldots, Y_n)$ for all $Y_1, \ldots, Y_n \in \mathcal{P}(E)$, then $\mathcal{F}(Q')(X_1, \ldots, X_n) = \mathcal{F}(Q)(X_1, \ldots, X_n)$ for all $X_1, \ldots, X_n \in \widetilde{\mathcal{P}}(E)$. This is an insensitivity property with respect to the exact choice of the domain as a whole. For example, we expect that the quantification result of "most tall are bald" does not depend on the precise choice of the domain as long as it is large enough to contain the fuzzy subsets **tall** and **bald** of interest. As has been shown in Glöckner (2000b, Theorem 16), this adequacy condition is satisfied by every DFS.

– *Multiplace quantification in $\mathcal{F}$ can always be reduced to one-place quantification* in $\mathcal{F}$ based on a suitably constructed quantifier and base set, see Glöckner (2000a, Theorem 35).

4.2 First example of an important DFS: $\mathcal{M}$

By stating the DFS axioms, we have made explicit our intuitions about "reasonable" mechanisms of fuzzy quantification. In order to show that these axioms are consistent (but also to make the theory useful for purposes of data fusion), we now present actual models. The model $\mathcal{M}$, to be described now, is the DFS that was used for the querying system mentioned in sec. 1 and it is the first DFS that was fully implemented for quantification. In Glöckner (1997), it was shown that $\mathcal{M}$ is a consistent generalisation of the fuzzification scheme proposed by Gaines (1976).
The model uses the fuzzy median as an aggregation operator over sets of gradual evaluations. The basic fuzzy median $\mathrm{med}_{\frac{1}{2}}$ is defined by

$$\mathrm{med}_{\frac{1}{2}}(u_1, u_2) = \begin{cases} \min(u_1, u_2) & : \quad \min(u_1, u_2) > \frac{1}{2} \\ \max(u_1, u_2) & : \quad \max(u_1, u_2) < \frac{1}{2} \\ \frac{1}{2} & : \quad \text{else} \end{cases}$$

for all $u_1, u_2 \in [0, 1]$, see Silvert (1979). The fuzzy median can be extended to an operator which accepts arbitrary subsets of $[0, 1]$ as its arguments. This (extended) fuzzy median is defined by

$$\mathrm{m}_{\frac{1}{2}} X = \mathrm{m}_{\frac{1}{2}}(\inf X,\ \sup X)\,,$$

for all $X \in \mathcal{P}([0, 1])$. In addition, we need the cut range $\mathcal{T}_\gamma(X) \subseteq \mathcal{P}(E)$ of a fuzzy subset X at the cutting level $\gamma \in [0, 1]$, which corresponds to a symmetrical, three-valued cut of X at γ:

$$\mathcal{T}_\gamma(X) = \{Y \subseteq E : X_\gamma^{\min} \subseteq Y \subseteq X_\gamma^{\max}\}$$

where

$$X_\gamma^{\min} = \begin{cases} X_{\geq \frac{1}{2}+\frac{1}{2}\gamma} & : \quad \gamma \in (0,\ 1] \\ X_{>\frac{1}{2}} & : \quad \gamma = 0 \end{cases}$$
$$X_\gamma^{\max} = \begin{cases} X_{>\frac{1}{2}-\frac{1}{2}\gamma} & : \quad \gamma \in (0,\ 1] \\ X_{\geq \frac{1}{2}} & : \quad \gamma = 0 \end{cases}$$

Here $X_{\geq\alpha} = \{e \in E : \mu_X(e) \geq \alpha\}$ denotes the α-cut, and $X_{>\alpha} = \{e \in E : \mu_X(e) > \alpha\}$ denotes the strict α-cut. γ can be thought of as a parameter of "cautiousness". If $\gamma = 0$, the set of indeterminates (i.e. of those $e \in E$ such that $e \in {X_i}_\gamma^{\max} \setminus {X_i}_\gamma^{\min}$) contains only those $e \in E$ with $\mu_{X_i}(e) = \frac{1}{2}$; all other elements of E are mapped to the closest truth value in $\{0,\ 1\}$. As γ increases, the set of indeterminates is increasing. For $\gamma = 1$, then, the level of maximal cautiousness is reached where all elements of E except those with $\mu_{X_i}(e) \in \{0,\ 1\}$ are interpreted as indeterminates.

By applying the extended fuzzy median to the quantification results obtained for all choices of arguments from the cut ranges, we are now able to interpret fuzzy quantifiers for any fixed choice of the cutting parameter. We hence define

$$Q_\gamma(X_1,\ \ldots,\ X_n) = \mathrm{m}_{\frac{1}{2}}\{Q(Y_1,\ \ldots,\ Y_n) : Y_1 \in \mathcal{T}_\gamma(X_1), \ldots, Y_n \in \mathcal{T}_\gamma(X_n)\}\,,$$

for all semi-fuzzy quantifiers $Q : \mathcal{P}(E)^n \longrightarrow [0,1]$ and all fuzzy arguments $X_1, \ldots, X_n$. This assignment $Q \mapsto Q_\gamma$ is not a DFS yet; the fuzzy median suppresses too much structure. We must hence take into account the results obtained at each level of cautiousness. This can be accomplished e.g. by means of integration, as is done in the following definition of the QFM $\mathcal{M}$. For every semi-fuzzy quantifier $Q : \mathcal{P}(E)^n \longrightarrow [0,1]$ and all fuzzy arguments $X_1, \ldots, X_n$, we define

$$\mathcal{M}(Q)(X_1,\ \ldots,\ X_n) = \int_0^1 Q_\gamma(X_1,\ \ldots,\ X_n)\, d\gamma\,.$$

Let us remark that the integral exists, regardless of Q and the choice of argument sets. It can be shown that $\mathcal{M}$ is a standard DFS. $\mathcal{M}$ is a practical model because it is both continuous in arguments and in quantifiers, i.e. robust against slight changes or noise in the fuzzy arguments $X_1, \ldots, X_n$ and in the definition of the quantifier Q. For the computation of $\mathcal{M}$ see sec. 4.5.

4.3 Second example of a DFS: $\mathcal{M}_{\mathbf{CX}}$

The integral, which was used in the definition of $\mathcal{M}$, is not the only possible way of abstracting from the cutting parameter γ. In Glöckner (2000a), the necessary and sufficient conditions have been spelled out that an aggregation mapping $\mathcal{B}$ needs to satisfy in order to make $\mathcal{M}_\mathcal{B}(Q)(X_1,\ \ldots,\ X_n) = \mathcal{B}((Q_\gamma(X_1,\ \ldots,\ X_n))_{\gamma\in[0,1]})$ a DFS. In addition, special properties of the resulting class of models, called $\mathcal{M}_\mathcal{B}$-DFSes, have been studied to some depth. In the course of this investigation, it turned out that there exists a model with unique adequacy properties. Even among the full class of standard models, it can be proven that $\mathcal{M}_{\mathrm{CX}}$ is the best possible model for quantification from a linguistic point of view because:

- it preserves convexity of absolute quantitative quantifiers like “about ten”, “between twenty and thirty”;
- it is the *only standard model* which permits the compositional interpretation of adjectival restriction by a fuzzy adjective, like in “almost all young A’s are B’s”. With $\mathcal{M}_{\mathrm{CX}}$, it is then guaranteed that $\mathcal{M}_{\mathrm{CX}}$(**almost all young**)$(A, B) = \mathcal{M}_{\mathrm{CX}}$(**almost all**)(**young** A, B);
- it is continuous in arguments and in quantifiers and hence robust against noise;
- it propagates fuzziness in arguments and in quantifiers, i.e. less specific input cannot result in more specific outputs;
- it combines easily with *finite* sets of truth values, i.e. no “new” truth values are generated as long as the finite subset of given truth values is closed under negation;
- is robust against consistent changes of membership grades, which is important because it is often hard to justify a specific choice of membership function.

$\mathcal{M}_{\mathrm{CX}}$ is hence the preferred choice for all applications that need to capture NL semantics (NLIs etc.). Here is its definition, which can be stated independently of the cut ranges and the median-based aggregation mechanism:

$$\mathcal{M}_{\mathrm{CX}}(Q)(X_1, \ldots, X_n) = \sup\{Q^L_{V,W}(X_1, \ldots, X_n) : V, W \in \mathcal{P}(E)^n, V_1 \subseteq W_1, \ldots, V_n \subseteq W_n\}$$

where

$$Q^L_{V,W}(X_1, \ldots, X_n) = \min(\Xi_{V,W}(X_1, \ldots, X_n), \inf\{Q(Y_1, \ldots, Y_n) : V_i \subseteq Y_i \subseteq W_i \text{ for all } i = 1, \ldots, n\})$$

$$\Xi_{V,W}(X_1, \ldots, X_n) \min_{i=1}^{n} \min(\inf\{\mu_{X_i}(e) : e \in V_i\}, \inf\{1 - \mu_{X_i}(e) : e \notin W_i\}) \,.$$

As has been remarked above, $\mathcal{M}_{\mathrm{CX}}$ is a standard DFS with unique adequacy properties. Another interesting aspect of $\mathcal{M}_{\mathrm{CX}}$ is that it consistently generalises the Sugeno integral and hence the "basic" FG-count approach, i.e. the well-behaved formula for $\mathrm{FG}^{(1)}_{\mathrm{abs}}$, to arbitrary n-place quantifiers, and to quantifiers that do not fulfill any special monotonicity requirements.

4.4 Third example of a DFS: $\mathcal{F}_{\mathbf{owa}}$

The class of known models has been further broadened broadened by abstracting from the median-based aggregation mechanism that was used to define $\mathcal{M}$ and (implicitely) $\mathcal{M}_{\mathrm{CX}}$. The use of $Q_\gamma(X_1, \ldots, X_n)$ in these models is then replaced with a pair of mappings which specify upper and lower bounds of the quantification results obtained for all choices of $Y_1, \ldots, Y_n$ in the cut ranges, viz

$$\top_{Q,X_1,\ldots,X_n}(\gamma) = \sup\{Q(Y_1, \ldots, Y_n) : Y_1 \in \mathcal{T}_\gamma(X_1), \ldots, Y_n \in \mathcal{T}_\gamma(X_n)\}$$
$$\bot_{Q,X_1,\ldots,X_n}(\gamma) = \inf\{Q(Y_1, \ldots, Y_n) : Y_1 \in \mathcal{T}_\gamma(X_1), \ldots, Y_n \in \mathcal{T}_\gamma(X_n)\} \,.$$

The cut ranges $\mathcal{T}_\gamma(X_i)$ are defined as in the case of $\mathcal{M}$. In Glöckner (2000c), the full class of models definable by $\mathcal{F}_\xi(Q)(X_1, \ldots, X_n) = \xi(\top_{Q,X_1,\ldots,X_n}, \bot_{Q,X_1,\ldots,X_n})$ has been investigated and the necessary and sufficient conditions on ξ have been presented which ensure that $\mathcal{F}_\xi$ be a DFS. As opposed to $\mathcal{M}_{\mathcal{B}}$-DFSes like $\mathcal{M}$ and $\mathcal{M}_{\mathrm{CX}}$, which are always guaranteed to propagate fuzziness in quantifiers and arguments, both conditions are optional for these more general DFSes. The last example $\mathcal{F}_{\mathrm{owa}}$, to be presented now, is representative of this new type of models. It is defined by

$$\mathcal{F}_{\mathrm{owa}}(Q)(X_1, \ldots, X_n) = \tfrac{1}{2}\int_0^1 \top_{Q,X_1,\ldots,X_n}(\gamma)d\gamma + \tfrac{1}{2}\int_0^1 \bot_{Q,X_1,\ldots,X_n}(\gamma)d\gamma \,.$$

$\mathcal{F}_{\mathrm{owa}}$ is a standard DFS. The model is of particular interest because it consistently generalises the Choquet integral and hence the "basic" OWA approach, i.e. the well-behaved formula for $\mathrm{OWA}^{(1)}_{\mathrm{prp}}$, to the "hard" cases of general multiplace and non-monotonic quantifiers. $\mathcal{F}_{\mathrm{owa}}$ is a practical model because it is continuous both in arguments and in quantifiers, which ensures a certain stability of the results against noise. In addition, the integral-based definition is suitable for developing gradient-descent training procedures. However, $\mathcal{F}_{\mathrm{owa}}$ does not propagate fuzziness in arguments nor in quantifiers. $\mathcal{F}_{\mathrm{owa}}$ is hence inferior to $\mathcal{M}_{\mathrm{CX}}$ from an adequacy perspective because less specific input can result in more specific output. Nevertheless, it can prove advantageous in applications where the inputs are overly fuzzy and a fine-grained result ranking is still needed, because $\mathcal{F}_{\mathrm{owa}}$ can discern cases in which all models that propagate fuzziness are stuck at the least specific result of $\frac{1}{2}$.

4.5 Computation of DFS quantifiers

Our theory has been designed to establish a principled account of fuzzy quantifiers in data fusion, and we have ensured this by stating it in the form of axioms. We have also provided specific models of the theory. Let us now discuss the computational aspects and show how the resulting quantifiers can be efficiently implemented.

Evaluation of "simple" quantifiers Let us assume some choice of domain $E \neq \varnothing$. The following definitions of two-valued "standard" quantifiers are straightforward:

$$\begin{aligned}
&\forall(Y) = 1 \Leftrightarrow Y = E\\
&\exists(Y) = 1 \Leftrightarrow Y \neq \varnothing\\
&\textbf{all}(Y_1,\, Y_2) = 1 \Leftrightarrow Y_1 \subseteq Y_2\\
&\textbf{some}(Y_1,\, Y_2) = 1 \Leftrightarrow Y_1 \cap Y_2 \neq \varnothing\\
&\textbf{no}(Y_1,\, Y_2) = 1 \Leftrightarrow Y_1 \cap Y_2 = \varnothing\\
&\textbf{at least k}(Y_1,\, Y_2) = 1 \Leftrightarrow |Y_1 \cap Y_2| \geq k\\
&\textbf{at most k}(Y_1,\, Y_2) = 1 \Leftrightarrow |Y_1 \cap Y_2| \leq k\\
&\textbf{more than k}(Y_1,\, Y_2) = 1 \Leftrightarrow |Y_1 \cap Y_2| > k\\
&\textbf{less than k}(Y_1,\, Y_2) = 1 \Leftrightarrow |Y_1 \cap Y_2| < k
\end{aligned}$$

In every standard DFS $\mathcal{F}$, these quantifiers are modeled as follows.

$$\begin{aligned}
&\mathcal{F}(\forall)(X) = \inf\{\mu_X(e) : e \in E\}\\
&\mathcal{F}(\exists)(X) = \sup\{\mu_X(e) : e \in E\}\\
&\mathcal{F}(\textbf{all})(X_1,\, X_2) = \inf\{\max(1 - \mu_{X_1}(e),\, \mu_{X_2}(e)) : e \in E\}\\
&\mathcal{F}(\textbf{some})(X_1,\, X_2) = \sup\{\min(\mu_{X_1}(e),\, \mu_{X_2}(e)) : e \in E\}\\
&\mathcal{F}(\textbf{no})(X_1,\, X_2) = \inf\{\max(1 - \mu_{X_1}(e),\, 1 - \mu_{X_2}(e)) : e \in E\}\\
&\mathcal{F}(\textbf{at least k})(X_1,\, X_2) = \sup\{\alpha \in [0,1] : |(X_1 \cap X_2)_{\geq\alpha}| \geq k\}\,,
\end{aligned}$$

for all $X,\ X_1,\ X_2 \in \widetilde{\mathcal{P}}(E)$. In particular, if E is finite, then

$$\mathcal{F}(\textbf{at least k})(X_1,\, X_2) = \mu_{[k]}(X_1 \cap X_2)$$

Note that **more than k** $=$ **at least k+1**, **less than k** $= 1 -$ **at least k** and **at most k** $= 1 -$ **more than k**, i.e. these quantifiers are covered by the formula for **at least k**.

Evaluation of quantitative one-place quantifiers We first need some observations on quantitative one-place quantifiers. We notice that a one-place semi-fuzzy quantifier $Q : \mathcal{P}(E) \longrightarrow [0,1]$ on a finite base set $E \neq \varnothing$ is quantitative if and only if there exists a mapping $\mu_{Q,E} : \{0, \ldots, |E|\} \longrightarrow [0,1]$ such that $Q(Y) = \mu_{Q,E}(|Y|)$, for all $Y \in \mathcal{P}(E)$. $\mu_{Q,E}$ is defined by

$$\mu_{Q,E}(j) = Q(Y_j)$$

for $j \in \{0, \ldots, |E|\}$, with $Y_j \in \mathcal{P}(E)$ an arbitrary subset of cardinality $|Y_j| = j$. Most natural language quantifiers satisfy a condition known as *having extension*, i.e. $E \subseteq E'$ entails that $Q_E(Y_1, \ldots, Y_n) = Q_{E'}(Y_1, \ldots, Y_n)$ for all $Y_1, \ldots, Y_n \in \mathcal{P}(E)$, where Q_E and $Q_{E'}$ are the models of Q in E and E', respectively. In this typical case that Q has extension, there exists $\mu_Q : \mathbb{N} \longrightarrow [0,1]$ such that for all finite base sets $E \neq \varnothing$, $\mu_{Q,E}(j) = \mu_Q(j)$ for all $j \in \{0, \ldots, |E|\}$.

Let us now simplify the formulas for $\bot_{Q,X_1,\ldots,X_n}(\gamma)$ and $\top_{Q,X_1,\ldots,X_n}\gamma$ in the case of quantitative Q, which is also useful for the median-based models because $Q_\gamma(X_1, \ldots, X_n) = \mathrm{med}_{\frac{1}{2}}(\top_{Q,X_1,\ldots,X_n}(\gamma), \bot_{Q,X_1,\ldots,X_n}(\gamma))$. Given a fuzzy subset $X \in \widetilde{\mathcal{P}}(E)$ of a finite base set $E \neq \varnothing$ and $\gamma \in [0,1]$, we abbreviate $|X|_\gamma^{\min} = |X_\gamma^{\min}|$ and $|X|_\gamma^{\max} = |X_\gamma^{\max}|$. For all $0 \leq \ell \leq u \leq |E|$, we further define

$$q^{\min}(\ell, u) = \min\{\mu_{Q,E}(k) : \ell \leq k \leq u\}$$
$$q^{\max}(\ell, u) = \max\{\mu_{Q,E}(k) : \ell \leq k \leq u\}\,.$$

For every quantitative one-place semi-fuzzy quantifier $Q : \mathcal{P}(E) \longrightarrow [0,1]$ on a finite base set, all $X \in \widetilde{\mathcal{P}}(E)$ and $\gamma \in [0,1]$, we then obtain that $\bot_{Q,X}(\gamma) = q^{\min}(\ell, u)$, $\top_{Q,X}(\gamma) = q^{\max}(\ell, u)$ and $Q_\gamma(X) = \mathrm{med}_{\frac{1}{2}}(q^{\min}(\ell, u), q^{\max}(\ell, u))$, abbreviating $\ell = |X|_\gamma^{\min}$ and $u = |X|_\gamma^{\max}$. In the frequent case that Q is *monotonic*, this reduces to

$$q^{\min}(\ell, u) = \mu_{Q,E}(\ell), \quad q^{\max}(\ell, u) = \mu_{Q,E}(u) \qquad \text{if } Q \text{ nondecreasing}$$
$$q^{\min}(\ell, u) = \mu_{Q,E}(u), \quad q^{\max}(\ell, u) = \mu_{Q,E}(\ell) \qquad \text{if } Q \text{ nonincreasing.}$$

The computation of convex (unimodal) quantifiers like "about 10" has been omitted here; it is discussed in Glöckner (2000b) and Glöckner and Knoll (2001).

In the case of the DFS $\mathcal{M}_{\mathrm{CX}}$, we can use the following *fuzzy interval cardinality* $\|X\|_{\mathrm{iv}} \in \widetilde{\mathcal{P}}(\mathbb{N} \times \mathbb{N})$ to evaluate quantitative one-place quantifiers, which is defined by

$$\mu_{\|X\|_{\mathrm{iv}}}(\ell, u) = \begin{cases} \min(\mu_{[\ell]}(X), 1 - \mu_{[u+1]}(X)) & : \quad \ell \leq u \\ 0 & : \quad \text{else} \end{cases} \qquad \text{for all } \ell, u \in \mathbb{N}.$$

Then for every quantitative one-place quantifier $Q : \mathcal{P}(E) \longrightarrow [0,1]$ on a finite base set and all $X \in \widetilde{\mathcal{P}}(E)$,

$$\mathcal{M}_{\mathrm{CX}}(Q)(X) = \max\{\min(\mu_{\|X\|_{\mathrm{iv}}}(\ell, u), q^{\min}(\ell, u)) : 0 \leq \ell \leq u \leq |E|\}\,.$$

It is instructive to notice that

$$\mu_{\text{FG-Count}(X)}(j) = \mu_{\|X\|_{\mathrm{iv}}}(j, |E|)\,, \qquad \mu_{\text{FE-Count}(X)}(j) = \mu_{\|X\|_{\mathrm{iv}}}(j, j+1)\,.$$

This explains why with general quantifiers μ_Q, the FG-count approach and the FE-count approach yield reasonable results in some cases (those where they coincide with $\mathcal{M}_{\mathrm{CX}}$) but fail in others.

For $\mathcal{M}$ and $\mathcal{F}_{\mathrm{owa}}$, a histogram-based approach can be used to efficiently implement the resulting quantifiers. For simplicity of presentation, we will describe a computation procedure suited to integer arithmetics. We hence assume that, for a fixed $m' \in \mathbb{N} \setminus \{0\}$, all membership values of fuzzy argument sets $X_1, \ldots, X_n$ satisfy

$$\mu_{X_i}(e) \in \left\{0, \frac{1}{m'}, \ldots, \frac{m'-1}{m'}, 1\right\} \qquad (*)$$

for all $e \in E$. If $X \in \widetilde{\mathcal{P}}(E)$ satisfies (*), then we can represent the required histogram of X as an $(m'+1)$-dimensional array $\mathrm{Hist}_X : \{0, \ldots, m'\} \longrightarrow \mathbb{N}$, defined by

$$\mathrm{Hist}_X[j] = \left|\{e \in E : \mu_X(e) = \frac{j}{m'}\}\right|$$

for all $j = 0, \ldots, m'$. We further assume that m' is even, (i.e. $m' = 2m$ for a given $m \in \mathbb{N} \setminus \{0\}$). The computation procedures for the DFSes $\mathcal{M}$ and $\mathcal{F}_{\mathrm{owa}}$ are presented in Table 3. In the algorithm for $\mathcal{M}$, we have utilized that $Q_\gamma(X) = \max(\frac{1}{2}, q^{\min}(\ell, u))$ if $Q_0(X) > \frac{1}{2}$ and $Q_\gamma(X) = \min(\frac{1}{2}, q^{\max}(\ell, u))$ otherwise. A further simplification is possible if Q is monotonic. For example, if Q is nondecreasing, then $q^{\min}(\ell, u) = \mu_{Q,E}(\ell)$ and $q^{\max}(\ell, u) = \mu_{Q,E}(u)$, i.e. we can omit the updating of u in the first for-loop and likewise omit ℓ in the second for-loop.

4.6 Evaluation of absolute quantifiers and quantifiers of exception

For every two-place semi-fuzzy quantifier $Q : \mathcal{P}(E)^2 \longrightarrow [0,1]$,

- Q is *absolute* iff there exists a quantitative one-place quantifier $Q' : \mathcal{P}(E) \longrightarrow [0,1]$ such that $Q(Y_1, Y_2) = Q'(X_1 \cap X_2)$ for all $Y_1, Y_2 \in \mathcal{P}(E)$.
- Q is called a *quantifier of exception* iff there exists an absolute quantifier $Q'' : \mathcal{P}(E)^2 \longrightarrow [0,1]$ such that Q is the antonym of Q'', i.e. $Q(Y_1, Y_2) = Q''(X_1, \neg X_2)$ for $Y_1, Y_2 \in \mathcal{P}(E)$.

For example, the two-place quantifier **about 50** is an absolute quantifier. Some examples of quantifiers of exception are presented in Table 4. The DFS axioms ensure that $\mathcal{F}(Q)(X_1, X_2) = \mathcal{F}(Q')(X_1 \cap X_2)$, whenever Q is an absolute quantifier and $X_1, X_2 \in \widetilde{\mathcal{P}}(E)$. Similarly if Q is a quantifier of exception, then $\mathcal{F}(Q)(X_1, X_2) = \mathcal{F}(Q')(X_1 \cap \neg X_2)$, for all $X_1, X_2 \in \widetilde{\mathcal{P}}(E)$, where $Q' : \mathcal{P}(E) \longrightarrow [0,1]$ is quantitative. We can hence use the algorithm for $\mathcal{F}(Q')(X)$, $\mathcal{F} \in \{\mathcal{M}_{\mathrm{CX}}, \mathcal{M}, \mathcal{F}_{\mathrm{owa}}\}$ to evaluate absolute quantifiers and quantifiers of exception.

Evaluation of proportional quantifiers A two-place semi-fuzzy quantifier $Q : \mathcal{P}(E)^2 \longrightarrow [0,1]$ on a finite base set is called *proportional* if there exist $v_0 \in [0,1]$, $f : [0,1] \longrightarrow [0,1]$ such that

$$Q(Y_1, Y_2) = \begin{cases} f(|Y_1 \cap Y_2|/|Y_1|) & : \quad Y_1 \neq \varnothing \\ v_0 & : \quad \text{else} \end{cases} \qquad \text{for all } Y_1, Y_2 \in \mathcal{P}(E).$$

Algorithm for computing $\mathcal{M}(Q)(X)$

```
INPUT: X
// initialise H, ℓ, u
H := Hist_X;
ℓ := Σ_{j=1}^{m} H[m+j];
u := ℓ + H[m];
cq := med_{1/2}(q^min(ℓ,u), q^max(ℓ,u));
if( cq == 1/2 ) return 1/2;
sum := cq;
if( cq > 1/2 )
    for( j := 1; j<m; j := j + 1 ) {
        nc:= true; // "no change"
        // update clauses for ℓ and u
        if( H[m+j] ≠ 0 )
            { ℓ := ℓ - H[m+j]; nc:= false; }
        if( H[m-j] ≠ 0 )
            { u := u + H[m-j]; nc:= false; }
        if( nc)
            { sum := sum + cq; continue; }
        // one of ℓ or u has changed
        cq := q^min(ℓ,u);
        if( cq ≤ 1/2 ) break;
        sum := sum + cq;
        }
else
    for( j := 1; j<m; j := j + 1 ) {
        nc:= true;
        .
        : // update clauses etc. as above
        // one of ℓ or u has changed
        cq := q^max(ℓ,u);
        if( cq ≥ 1/2 ) break;
        sum := sum + cq;
        }
return (sum + 1/2*(m-j)) / m;
END
```

Algorithm for computing $\mathcal{F}_{\text{owa}}(Q)(X)$

```
INPUT: X
// initialise H, ℓ, u
H := Hist_X;
ℓ := Σ_{j=1}^{m} H[m+j];
u := ℓ + H[m];
cq := q^min(ℓ,u) + q^max(ℓ,u);
sum := cq;
for( j := 1; j<m; j := j + 1 ) {
    ch := false; // "change"
    // update clauses for ℓ and u
    if( H[m+j] ≠ 0 )
        { ℓ := ℓ - H[m+j]; ch := true; }
    if( H[m-j] ≠ 0 )
        { u := u + H[m-j]; ch := true; }
    if( ch )
        // one of ℓ or u has changed
        { cq := q^min(ℓ,u) + q^max(ℓ,u); }
    sum := sum + cq;
    }
return sum / m'; // where m' = 2*m
END
```

Table3. Algorithms for evaluating quantitative one-place quantifiers

Quantifier	Antonym (absolute)
all	**no**
all except exactly k	**exactly k**
all except about k	**about k**
all except at most k	**at most k**

Table4. Examples of quantifiers of exception

For example, we can provide a definition of **almost all** by choosing $f(z) = S(x, 0.7, 0.9)$ and $v_0 = 1$, where S is Zadeh's S-function. Usually f and v_0 can be chosen independently of E, i.e. Q has extension. We shall restrict our attention to those proportional quantifiers where $f : [0,1] \longrightarrow [0,1]$ is nondecreasing. (If f is nonincreasing, we can compute $\mathcal{F}(Q) = \neg\mathcal{F}(\neg Q)$, noting that the negation $\neg Q$ is proportional and nondecreasing.) Suppose Q is such a quantifier and $X_1, X_2 \in \widetilde{\mathcal{P}}(E)$. We are using abbreviations $Z_1 = X_1$, $Z_2 = X_1 \cap X_2$ and $Z_3 = X_1 \cap \neg X_2$; further let $\ell_k = |Z_k|_\gamma^{\min}$ and $u_k = |Z_k|_\gamma^{\max}$, $k \in \{1,2,3\}$, $f^{\min} = f(\ell_2/(\ell_2 + u_3))$ and $f^{\max} = f(u_2/(u_2 + \ell_3))$. Then

$$Q_\gamma(X_1, X_2) = \text{med}_{\frac{1}{2}}(q^{\min}(\ell_1, \ell_2, u_1, u_3),\, q^{\max}(\ell_1, \ell_3, u_1, u_2))\,,$$

where the definitions of $q^{\min}, q^{\max} : \{0, \ldots, |E|\}^4 \longrightarrow [0, 1]$ are as follows.

1. $\ell_1 > 0$. Then $q^{\min} = f^{\min}$.
2. $\ell_1 = 0$.
   a. $\ell_2 + u_3 > 0$.
      Then $q^{\min} = \min(v_0,\ f^{\min})$.
   b. $\ell_2 + u_3 = 0$.
      i. $u_1 > 0$.
         Then $q^{\min} = \min(v_0, f(1))$.
      ii. $u_1 = 0$. Then $q^{\min} = v_0$.

Note. If $v_0 \leq f(1)$, then $\min(v_0, f(1)) = v_0$, i.e. we need not distinguish 2.b.i and 2.b.ii.
For $q^{\max}(\ell_1, \ell_3, u_1, u_2)$, we have:

1. $\ell_1 > 0$. Then $q^{\max} = f^{\max}$.
2. $\ell_1 = 0$.
   a. $u_2 + \ell_3 > 0$.
      Then $q^{\max} = \max(v_0,\ f^{\max})$.
   b. $u_2 + \ell_3 = 0$.
      i. $u_1 > 0$.
         Then $q^{\max} = \max(v_0,\ f(0))$.
      ii. $u_1 = 0$. Then $q^{\max} = v_0$.

Note. If $f(0) \leq v_0$, then 2.b.i and 2.b.ii need not be distinguished.

The algorithms for evaluating proportional quantifiers are presented in table 5. As shown in part a. of the table, a slight modification of the algorithm for $\mathcal{M}(Q)(X)$ in the case of one-place quantitative quantifiers is sufficient to compute $\mathcal{M}(Q)(X_1, X_2)$ for proportional quantifiers. Part b. of the table depicts the algorithm for evaluating $\mathcal{M}_{\mathrm{CX}}(Q)(X_1, X_2)$ in proportional case. The algorithm for computing $\mathcal{F}_{\mathrm{owa}}(Q)(X)$ can be adapted in a similar fashion to implement $\mathcal{F}_{\mathrm{owa}}(Q)(X_1, X_2)$ for proportional Q.

5 Results of applying DFS to image data

In this section we demonstrate the behaviour of DFS quantification when applied to our standard image data set. These results are based on the DFS $\mathcal{M}$, which was the first model to be implemented and applied in the multimedia retrieval system for weather documents. In the first example, the criterion for evaluating $\mathcal{M}$ is "As much as possible of Southern Germany is cloudy". This criterion is reduced to a corresponding condition on discrete pixels in the digitized image, viz "As many pixels as possible that belong to Southern Germany are classified as cloudy".

The ranking was hence computed by evaluating $\mathcal{M}$(**as many as possible**), where **as many as possible**$(X, Y) = |X \cap Y|/|X|$. As desired, the sequence of images (from left to right, from top to bottom) shows a clear decrease of cloudiness in Southern Germany.

In the next example we demonstrate the suitability of fuzzy quantifiers to aggregate and fuse data, see Fig. 10. The basic domain is a set of point in time. Over every image pixel, two fuzzy sets are defined: (a) the membership grade of the point in time to the fuzzy temporal condition "in the last days", and (b) the degree of membership to which the pixel is classified as "cloudy"

a. Algorithm for computing $\mathcal{M}(Q)(X_1, X_2)$

```
INPUT: X_1, X_2
// initialise H_k, ℓ, u
H_1 := Hist_{X_1};
H_2 := Hist_{X_1 ∩ X_2};
H_3 := Hist_{X_1 ∩ ¬X_2};
for( k := 1; k ≤ 3; k := k+1 ) {
    ℓ_k := Σ_{j=1}^{m} H_k[m+j]; u_k := ℓ_k + H_k[m];
    }
cq := med_{1/2}(q^min(ℓ_1, ℓ_2, u_1, u_3), q^max(ℓ_1, ℓ_3, u_1, u_2))
if( cq == 1/2 ) return 1/2;
sum := cq;
if( cq > 1/2 )
    for( j := 1; j<m; j := j + 1 ) {
        nc:= true; // "no change"
        // update clauses for ℓ_1, ℓ_2, u_1, u_3
        if( H_1[m+j] ≠ 0 )
            { ℓ_1 := ℓ_1 - H_1[m+j]; nc:= false; }
        if( H_2[m+j] ≠ 0 )
            { ℓ_2 := ℓ_2 - H_2[m+j]; nc:= false; }
        if( H_1[m-j] ≠ 0 )
            { u_1 := u_1 + H_1[m-j]; nc:= false; }
        if( H_3[m-j] ≠ 0 )
            { u_3 := u_3 + H_3[m-j]; nc:= false; }
        if( nc)
            { sum := sum + cq; continue; }
        // one of ℓ_1, ℓ_2, u_1, u_3 has changed
        cq := q^min(ℓ_1, ℓ_2, u_1, u_3);
        if( cq ≤ 1/2 ) break;
        sum := sum + cq;
        }
else
    for( j := 1; j<m; j := j + 1 ) {
        nc:= true;
        // update clauses for ℓ_1, ℓ_3, u_1, u_2
        if( H_1[m+j] ≠ 0 )
            { ℓ_1 := ℓ_1 - H_1[m+j]; nc:= false; }
        if( H_3[m+j] ≠ 0 )
            { ℓ_3 := ℓ_3 - H_3[m+j]; nc:= false; }
        if( H_1[m-j] ≠ 0 )
            { u_1 := u_1 + H_1[m-j]; nc:= false; }
        if( H_2[m-j] ≠ 0 )
            { u_2 := u_2 + H_2[m-j]; nc:= false; }
        if( nc)
            { sum := sum + cq; continue; }
        // one of ℓ_1, ℓ_3, u_1, u_2 has changed
        cq := q^max(ℓ_1, ℓ_3, u_1, u_2);
        if( cq ≥ 1/2 ) break;
        sum := sum + cq;
        }
return (sum + 1/2*(m-j)) / m;
END
```

b. Algorithm for computing $\mathcal{M}_{\mathrm{CX}}(Q)(X_1, X_2)$

```
INPUT: X_1, X_2
// initialise H_k, ℓ, u
H_1 := Hist_{X_1};
H_2 := Hist_{X_1 ∩ X_2};
H_3 := Hist_{X_1 ∩ ¬X_2};
for( k := 1; k ≤ 3; k := k+1 ) {
    ℓ_k := Σ_{j=1}^{m} H_k[m+j]; u_k := ℓ_k + H_k[m];
    }
cq := med_{1/2}(q^min(ℓ_1, ℓ_2, u_1, u_3), q^max(ℓ_1, ℓ_3, u_1, u_2))
if( cq == 1/2 ) return 1/2;
sum := cq;
if( cq > 1/2 )
{
    for( j := 1; j<m; j := j + 1 ) {
        ch := false; // "change"
        // update clauses for ℓ_1, ℓ_2, u_1, u_3
        if( H_1[m+j] ≠ 0 )
            { ℓ_1 := ℓ_1 - H_1[m+j]; ch := true; }
        if( H_2[m+j] ≠ 0 )
            { ℓ_2 := ℓ_2 - H_2[m+j]; ch := true; }
        if( H_1[m-j] ≠ 0 )
            { u_1 := u_1 + H_1[m-j]; ch := true; }
        if( H_3[m-j] ≠ 0 )
            { u_3 := u_3 + H_3[m-j]; ch := true; }
        if( ch )
            // one of ℓ_1, ℓ_2, u_1, u_3 has changed
            { cq := q^min(ℓ_1, ℓ_2, u_1, u_3); }
        if( cq ≤ m+j )
            { return (m+j)/m′ }
        }
    return 1;
}
else
    for( j := 1; j<m; j := j + 1 ) {
        ch := false;
        // update clauses for ℓ_1, ℓ_3, u_1, u_2
        if( H_1[m+j] ≠ 0 )
            { ℓ_1 := ℓ_1 - H_1[m+j]; ch := true; }
        if( H_3[m+j] ≠ 0 )
            { ℓ_3 := ℓ_3 - H_3[m+j]; ch := true; }
        if( H_1[m-j] ≠ 0 )
            { u_1 := u_1 + H_1[m-j]; ch := true; }
        if( H_2[m-j] ≠ 0 )
            { u_2 := u_2 + H_2[m-j]; ch := true; }
        if( ch )
            // one of ℓ_1, ℓ_3, u_1, u_2 has changed
            { cq := q^max(ℓ_1, ℓ_3, u_1, u_2); }
        if( cq ≥ m-j )
            { return (m-j)/m′; }
        }
return 0;
END
```

Table5. Algorithms for evaluating two-place proportional quantifiers

(at the respective point in time). These two fuzzy sets are then aggregated by applying the fuzzy quantifiers, and the result is a grey scale image sequence under the given aggration criterion. It can be seen in the image sequence that in all regions there have been clouds over the days and that in some of these regions it was always cloudy. Regions that meet the criterion are in white, regions that do not meet with the criterion are in black. To compute the results, the model $\mathcal{M}$ was used and the quantifier trp (for trapezoidal) that was applied is given by the following formula:

$$\mathbf{trp}_{a,b,c}(Y_1, Y_2) = \begin{cases} t_{a,b}(|Y_1 \cap Y_2|/|Y_1|) & : \quad Y_1 \neq \varnothing \\ c & : \quad Y_1 = \varnothing \end{cases} \qquad t_{a,b}(z) = \begin{cases} 0 & : \quad z < a \\ \frac{z-a}{b-a} & : \quad a \leq z \leq b \\ 1 & : \quad z > b \end{cases}$$

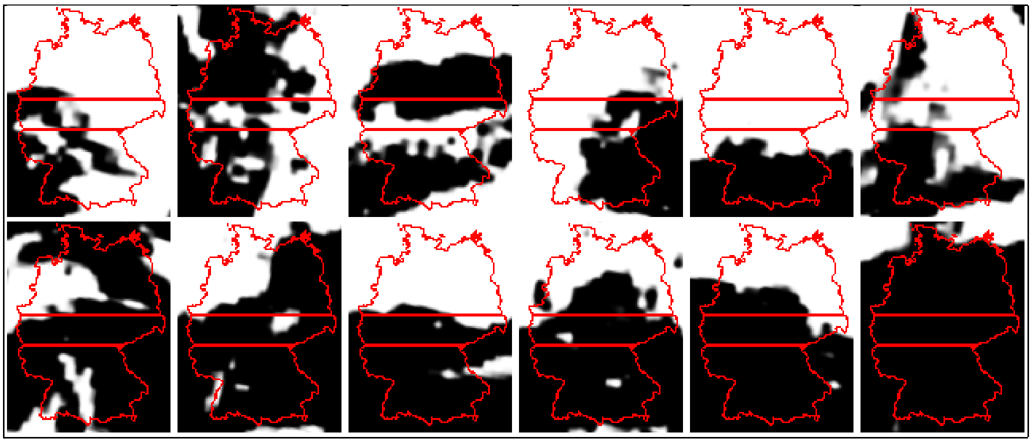

Figure9. Ranking computed for query “**As much as possible of Southern Germany is cloudy**”

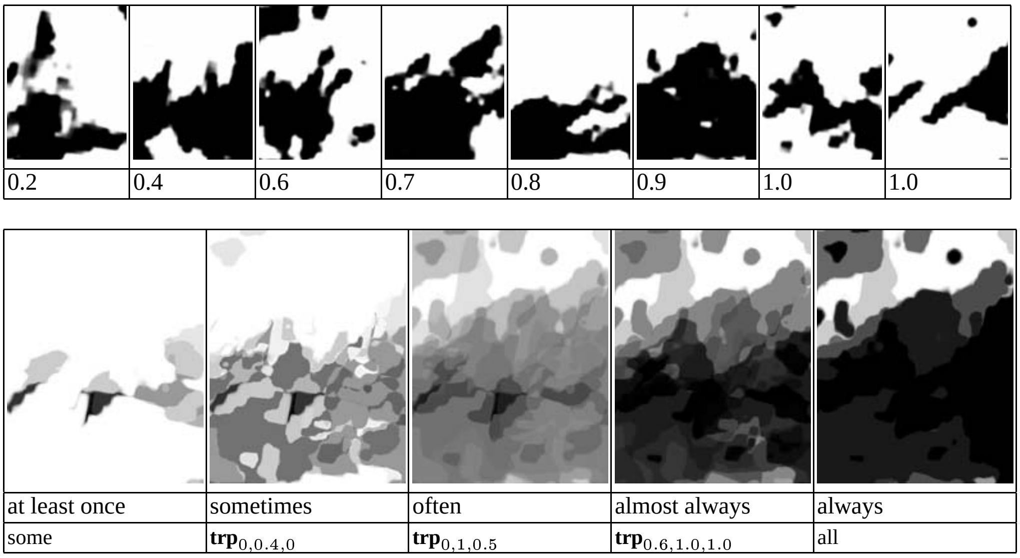

Figure10. Image sequence and fusion results for various choices of the criterion “**Q-times cloudy in the last days**”.

6 Conclusions

Fuzzy quantifiers are, in principle, a powerful method for combining and summarising information. As they are based on a linguistic foundation, they are easy to use and to understand for humans. In particular, it is much easier to incorporate expert knowledge into the fusion process by means of linguistic descriptions than it is, for example, to construct statistical models. Morevover, the linguistic approach also holds the potential to specify "fusion plans", i.e. *what* is to be fused *when* and *how* (see Dasarathy and Townsend (1999)). All of the existing approaches to fuzzy quantification introduced so far, however, give rise to implausible results in the most relevant situations. From a linguistic point of view, none of these approaches is suited to model two-place quantification with proportional quantifiers, which is the most important case. The reason for their poor performance is that they ignore important results from linguistics. To improve fuzzy quantification, methods from both fuzzy set theory and linguistics must be combined. The most useful tool for this undertaking we found to be the theory of generalised quantifiers, TGQ.

The theory of fuzzy quantification we have developed on this linguistic groundwork, DFS, improves upon existing approaches because it consistently extends the linguistic theory of NL quantification (i.e., TGQ) to incorporate gradual quantification results and fuzzy argument sets. Its axiomatic foundation guarantees conformance with linguistic adequacy considerations. Some of the models consistently generalize existing approaches to the hard cases of non-monotonic and multiplace quantification ($n \geq 2$); i.e. these models generalize the basic FG-count approach/Sugeno integral and the basic OWA approach/Choquet integral. We have also developed an efficient implementation based on histogram computations. Because of its axiomatic foundation, all models of DFS theory (i.e. all DFSes) are guaranteed to exhibit the desired adequacy properties.

Therefore, for the first time, a linguistically sound computational theory of fuzzy quantification is available, which is a precondition of convincing applications to linguistic data fusion.

References

J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.

J. van Benthem. Determiners and logic. *Linguistics and Philosophy*, 6, 1983.

G. Bordogna and G. Pasi. A fuzzy information retrieval system handling users' preferences on document sections. In D. Dubois, H. Prade, and R.R. Yager, editors, *Fuzzy Information Engineering*. Wiley, New York, 1997.

R. Brooks and S. Iyengar. *Multi-Sensor Fusion: Fundamentals and Applications with Software*. Prentice-Hall, 1997.

B. Dasarathy and S. Townsend. FUSE – fusion utility sequence estimator. In *Proc. 2nd Int. Conf. on Information Fusion*, Sunnyvale, CA, 1999. Int. Society of Information Fusion.

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

I. Glöckner and A. Knoll. Application of fuzzy quantifiers in image processing: A case study. In L.C. Jain, editor, *Proceedings of the Third International Conference on Knowledge-Based Intelligent Information Engineering Systems (KES '99)*, pages 259–262, Adelaide, Australia, Aug./Sep. 1999. IEEE.

I. Glöckner and A. Knoll. Natural-language navigation in multimedia archives: An integrated approach. In *Proceedings of the Seventh ACM Multimedia Conference (MM '99)*,, pages 313–322, Orlando, Florida, Oct./Nov. 1999.

I. Glöckner and A. Knoll. Architecture and retrieval methods of a search assistant for scientific libraries. In R. Decker and W. Gaul, editors, *Classification and Information Processing at the Turn of the Millennium*, pages 460–468. Springer, Heidelberg, 2000.

I. Glöckner and A. Knoll. A formal theory of fuzzy natural language quantification and its role in granular computing. In W. Pedrycz, editor, *Granular Computing: An Emerging Paradigm*. Physica-Verlag, 2001. to appear.

I. Glöckner. DFS – an axiomatic approach to fuzzy quantification. TR97-06, Technische Fakultät, Universität Bielefeld, P.O.-Box 100131, 33501 Bielefeld, Germany, 1997.

I. Glöckner. A framework for evaluating approaches to fuzzy quantification. TR99-03, Technische Facultät, Universität Bielefeld, P.O.-Box 100131, 33501 Bielefeld, Germany, 1999.

I. Glöckner. Advances in DFS theory. TR2000-01, Technische Fakultät, Universität Bielefeld, P.O.-Box 100131, 33501 Bielefeld, Germany, 2000.

I. Glöckner. An axiomatic theory of fuzzy quantifiers in natural languages. TR2000-03, Technische Fakultät, Universität Bielefeld, P.O.-Box 100131, 33501 Bielefeld, Germany, 2000.

I. Glöckner. A broad class of standard DFSes. TR2000-02, Technische Fakultät, Universität Bielefeld, P.O.-Box 100131, 33501 Bielefeld, Germany, 2000.

I. Goodman, R. Mahler, and H. Nguyen. Mathematics of data fusion. In *Theory and Decision Library. Series B, Mathematical and Statistical Methods*, volume 37. Kluwer Academic, 1997.

D. Hall. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, 1992.

E. Keenan and J. Stavi. A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9, 1986.

A. Knoll, C. Altenschmidt, J. Biskup, H.-M. Blüthgen, I. Glöckner, S. Hartrumpf, H. Helbig, C. Henning, Y. Karabulut, R. Lüling, B. Monien, T. Noll, and N. Sensen. An integrated approach to semantic evaluation and content-based retrieval of multimedia documents. In C. Nikolaou and C. Stephanidis, editors, *Research and Advanced Technology for Digital Libraries: Proceedings of ECDL '98*, pages 409–428. Springer, Berlin, New York, 1998.

A.L. Ralescu. A note on rule representation in expert systems. *Information Sciences*, 38:193–203, 1986.

D. Ralescu. Cardinality, quantifiers, and the aggregation of fuzzy criteria. *Fuzzy Sets and Systems*, 69, 1995.

W. Silvert. Symmetric summation: A class of operations on fuzzy sets. *IEEE Transactions on Systems, Man, and Cybernetics*, 9:657–659, 1979.

H. Thiele. On T-quantifiers and S-quantifiers. In *The Twenty-Fourth International Symposium on Multiple-Valued Logic*, pages 264–269, Boston, MA, 1994.

W.G. Waller and D.H. Kraft. A mathematical model of a weighted boolean retrieval system. *Information Processing & Management*, 15:235–245, 1979.

R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics*, 18(1):183–190, Jan./Feb. 1988.

R.R. Yager. Connectives and quantifiers in fuzzy sets. *Fuzzy Sets and Systems*, 40:39–75, 1991.

R.R. Yager. Counting the number of classes in a fuzzy set. *IEEE Trans. on Systems, Man, and Cybernetics*, 23(1):257–264, 1993.

L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8,9:199–249,301–357, 1975.

L.A. Zadeh. A theory of approximate reasoning. In J. Hayes, D. Michie, and L. Mikulich, editors, *Machine Intelligence*, volume 9, pages 149–194. Halstead, New York, 1979.

L.A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Computers and Mathematics with Applications*, 9:149–184, 1983.