Universität
Bielefeld

Forschungsberichte der
Technischen Fakultät
Abteilung Informationstechnik

Fuzzy Quantifiers for Processing Natural Language Queries in Content-Based Multimedia Retrieval Systems

Ingo Glöckner and Alois Knoll

Report 97-05

Universität Bielefeld • Postfach 10 01 31 • 33501 Bielefeld • FRG

Impressum: Herausgeber:
Robert Giegerich, Alois Knoll, Helge Ritter, Gerhard Sagerer,
Ipke Wachsmuth

Anschrift:
Technische Fakultät der Universität Bielefeld,
Abteilung Informationstechnik, Postfach 10 01 31, 33501 Bielefeld

ISSN 0946-7831

Contents

Abstract

The processing of natural language (NL) queries and the search for semantic matches between such queries and the contents of multimedia documents necessitate powerful quantifiers that adequately model quantifying expressions in NL.

In this report, we develop an axiomatic theory of fuzzy NL quantifiers and present a model of this theory which makes its systematical interpretation possible. The need for such a theory arises from the fact that the meaning of a natural language query depends heavily not only on the concepts it contains, but also on the various quantifying expressions interrelating these concepts in the query, which are often fuzzy in nature.

The resulting operators form a class of generic operators for information aggregation and the fusion of gradual evaluations, which could also prove useful in more traditional retrieval systems.

We conclude the report by sketching some applications of the theory to our multimedia retrieval system and give an example of how it is used for the content-based retrieval of meteorological documents.[1]

Keywords

natural-language query processing, content-based image retrieval, multimedia systems, fuzzy information retrieval, fuzzy quantifiers, generalized quantifiers, information aggregation.

[1]Parts of the work reported here were funded by the ministry of science and research of the state of Northrhine-Westphalia within the collaborative research initiative "Virtual Knowledge Factory".

1 Introduction

1.1 Research background

The theory of fuzzy generalized quantification presented here has evolved from our research into the design of a semantic search module for the multimedia retrieval system HPQS (High-Performance Query Server) [1], a system for structuring and querying large databases of multimedia documents, which is accessed by natural language (NL) queries.

HPQS attempts to cover the full path from natural language input to the presentation of retrieved documents. Queries are formulated in natural language and are evaluated for their semantic contents. For the document evaluation, a knowledge model consisting of a set of domain specific concept interpretation methods is constructed. Thus, the semantics of both the query and the documents can be interconnected, i.e. the retrieval process searches for a match on the *semantic level* (not merely on the level of keywords or global image properties) between the query and the document. Methods from fuzzy set theory are used to find the matches. Furthermore, the retrieval methods associate information from different document classes (texts, images, …). To avoid the loss of information inherent to pre-indexing, documents need not be indexed; in principle, every search may be performed on the raw data under a given query. The system can therefore answer every query that can be expressed in the semantic model. To achieve the high data rates necessary for on-line analysis, dedicated VLSI search processors are being developed along with a parallel high-throughput media-server. A further speed-up is achieved by the mediator module [2] which, apart from hiding various kinds of implementation details, maintains an intelligent result cache. This caching mechanism plays a similar role in HPQS as that of pre-computed indexes in more traditional retrieval systems. The system architecture is sketched in fig. 1.

A prototype implementation, for which the domain of meteorology was chosen, serves to validate the HPQS architecture and the content-based search mechanism; the document base of the HPQS system contains large amounts of textual and multimedia weather reports (short, mid and long-term), weather maps, satellite images and time-series.

The range of queries is large, starting from simple requests for documents about the current weather, by the way of queries about weather developments in the past fulfilling certain constraints, to questions about special weather conditions in certain locations. Sample queries that can be answered are:

- *Show me a map of the region with the hottest temperature readings yesterday!*
- *What additional weather data do you have about that region?*
- *Show me weather maps with a similar temperature distribution!*
- *Can you show me a satellite image of upper Bavaria that has a 50% coverage of clouds?*

1.2 The case for an NL interface

The acceptance of a multimedia system depends crucially on the design of its user interface. Ideally, the user interface should totally hide the complexity of the program, thus providing the view of an easy-to-use "information assistant". In particular, it should meet the needs of naive users who may be competent specialists in their field, but do not know or do not want to learn the peculiarities

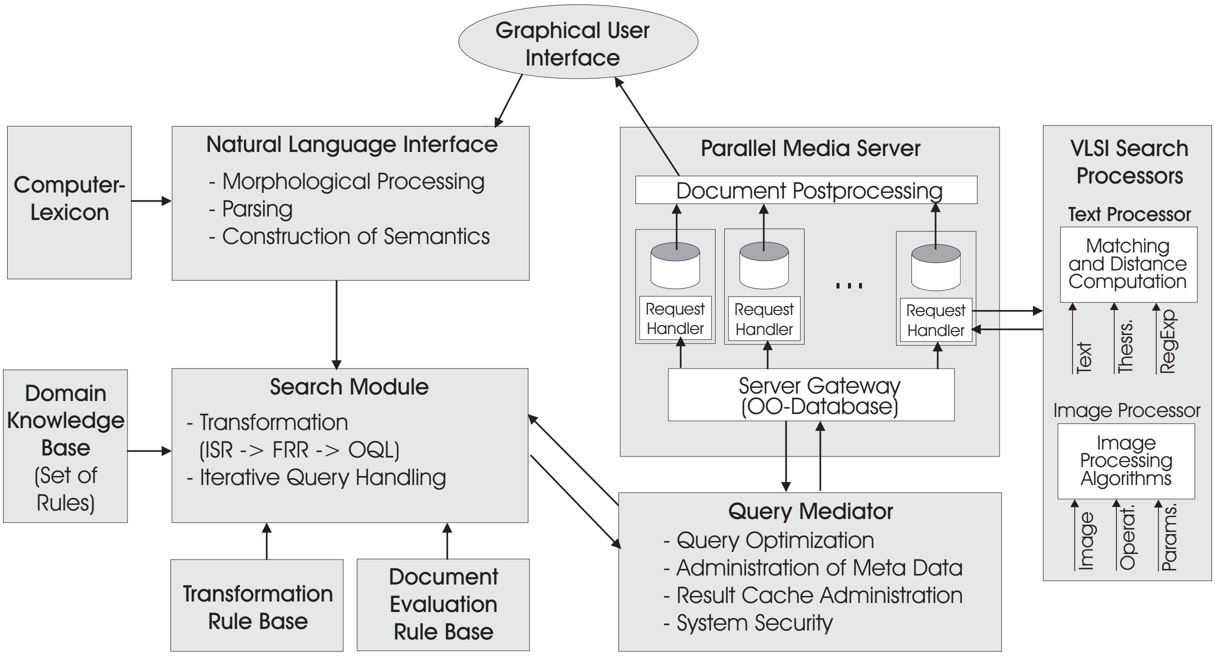

Figure 1: Architecture of the HPQS

of retrieval techniques or languages. Natural language provides an *intuitive interface* because everybody knows how to use their native language without effort. Hence, providing an NL front-end can remove technical barriers in accessing the more advanced features of an information retrieval system. Moreover, a reliable NL front end is a precondition for future retrieval systems that are controlled by speech input.

With non-textual documents, e.g. still images, graphics, or technical drawings, there is an even stronger case for natural language interfaces. While a text normally has a central "theme" that can be distilled into keywords, pictures may contain a multitude of different objects. It is not only these objects but often their relations that constitute the interesting part of the image content ("a car *in front* of the house"). Quantitative aspects are best expressed using quantifiers ("*several* cars in front of the house").

Relational and quantitative information also plays a crucial role in the domain of meteorology. It makes a difference to users planning their vacation in Italy whether "there are clouds" on the current weather image or whether "there are clouds in Italy". The difference in meaning is easily expressed in natural language. Keyword-based approaches, however, are obviously not suited for querying such structured information.

1.3 Retrieval methods for NL queries

Querying by natural language is highly demanding not only because the NL-analysis itself is a complex endeavour, but also because the users express their information needs on the level of *meaning*, the processing of which requires content-based search. For example, when a query requests images displaying cloudiness in some geographic region, then a corresponding interpretation method must be available in the system which evaluates images of the document base for their "cloudiness in a specified region". Natural language concepts such as "cloudy", "warm" and "fine weather", however, are inherently fuzzy. For example, there is no sharp-cut threshold in cloud density which distinguishes a "cloudy" weather situation from a "non-cloudy" one.[2] Therefore, we decided to model these concepts by fuzzy evaluation methods which result in gradual evaluations. Although these methods build on a common substrate of generic image processing methods, they are highly domain specific. An example is given at the end of the paper.

The fuzziness of natural language queries is also observed in local specifications (e.g. "upper Bavaria", "near Munich") and temporal specifications ("recently", "about one month ago"), which are accounted for in our system. These are modelled as fuzzy regions in time or space.

1.4 Quantifiers in NL queries

In traditional information retrieval systems, the set of operators for the aggregation of search results is essentially restricted to the Boolean connectives "and", "or" and "not".[3] These connectives may, of course, also occur in NL queries, and they form an integral part of the meaning of these queries. For example, when computing search results for an NL query containing a condition "not cloudy", the connective "not" must not be ignored.

The "modes of combination" of natural language, however, i.e. the various ways in which concepts might be interrelated, are by no means restricted to these connectives. In particular, the meaning of NL queries depends heavily on the *quantifying expressions* involved, as e.g. witnessed by the different meaning of "there are few clouds over Italy" vs. "there is a lot of clouds over Italy", which both could be part of queries addressed to our system. Therefore, the meaning of quantifying expressions must be accounted for if content-based retrieval, which reflects the semantics of NL queries, is to be achieved.

Let us note that many of these expressions are *approximate* or *fuzzy* in nature:

- $\mathbf{often}$, $\mathbf{rarely}$, $\mathbf{recently}$, $\mathbf{mostly}$, $\mathbf{almost\ always}$,... (temporal)
- $\mathbf{almost\ everywhere}$, $\mathbf{hardly\ anywhere}$, $\mathbf{partly}$,... (local)
- $\mathbf{many}$, $\mathbf{few}$, $\mathbf{a\ few}$, $\mathbf{almost\ all}$, $\mathbf{about\ ten}$, $\mathbf{about\ 40\ percent}$,... (approximate specification of the cardinality of a set, or a proportion of cardinalities).

These expressions are best modelled as resulting in gradual evaluations.

Another class of quantificational expressions is not itself fuzzy, but an application of these quantifiers to fuzzy concepts presupposes a reasonable and systematic generalisation of their semantics to the case of fuzzy arguments. For example, $\mathbf{everywhere}$, $\mathbf{nowhere}$, $\mathbf{always}$, $\mathbf{ten\ times}$, $\mathbf{at\ least\ ten}$, $\mathbf{all}$, $\mathbf{less\ than\ 20}$,... can be adequately modelled in the framework of a classical (two-valued) logic. The query, however,

Is the weather fine in all of upper Bavaria?

requires the two-valued quantifier $\mathbf{all}$ to be applied to the fuzzy regions $\mathbf{upper\ Bavaria}$, $\mathbf{fine\ weather}$. In order to do so, the semantics of $\mathbf{all}$ must be extended to fuzzy arguments.

[2] At least in the non-technical use of the word "cloudy" which we intend to model.

[3] Adjacency ("near") is not an aggregation operator because it applies to search terms instead of logical evaluations.

In the sequel, a general theory based on axioms describing the intended behaviour of fuzzy quantifiers is presented which permits a *systematic* interpretation of natural language queries involving fuzzy quantifiers.

Finally, we sketch some applications of the theory to multimedia retrieval and give an example of how it is used in a system for the content-based retrieval of meteorological documents.

Since in colloquial language, people are using natural language quantifiers with the same ease as the connectives “and”, “or”, and “not”, these quantifiers constitute a rich yet *human understandable* class of operators for information aggregation and data fusion. For example, they can be used in textual information retrieval—in a similar fashion as the operators “and” and “or” in traditional query languages—to provide for more natural (e.g., compromising) ways of aggregating over sets of search terms. Availibility of these NL quantifiers makes possible a more fine-grained and purposive search compared to standard methods.

As opposed to the domain-dependent evaluation methods, our approach to fuzzy quantifiers is completely domain independent, i.e. it will prove useful in very different applications both

- for handling the fuzziness and vagueness of natural language expressions and
- for associating or fusing information from different document sources, evaluation methods, or even from different domains.

2 A theory of fuzzy generalized quantification

2.1 Determiners

Our solution to the problem of systematic interpretation of fuzzy quantifying expressions is based on the Theory of Generalized Quantifiers (TGQ) [3, 4, 5, 6], certainly the most elaborate extensional theory of natural language quantification. In TGQ, quantifying expressions are generally called “determiners”.

Definition 1 (Determiner) *An n-ary determiner on a set of “entities” or “individuals” E is a function $D : \mathcal{P}(E)^n \longrightarrow 2$, where $2 = \{0, 1\}$ is the set of truth values and $\mathcal{P}(E)$ is the set of subsets of E.*

A determiner D thus assigns to each n-tuple $(X_1, \ldots, X_n)$ of subsets of E a corresponding truth value $D(X_1, \ldots, X_n) \in 2$. E might be any set of objects; in our protoype system, a set of time points, a set of pixel coordinates, a set of image regions, or a set of text search expressions. Well-known examples of determiners are

$$\begin{aligned}
\forall(X) = 1 &\quad\Leftrightarrow\quad X = E \\
\exists(X) = 1 &\quad\Leftrightarrow\quad X \neq \varnothing \\
\mathbf{all}(X_1, X_2) = 1 &\quad\Leftrightarrow\quad X_1 \subseteq X_2 \\
\mathbf{some}(X_1, X_2) = 1 &\quad\Leftrightarrow\quad X_1 \cap X_2 \neq \emptyset \\
\mathbf{atleast\ n}(X_1, X_2) = 1 &\quad\Leftrightarrow\quad \mathrm{card}\, X_1 \cap X_2 \geq n
\end{aligned}$$

So if an image J contains objects $E = \{c_1, c_2\}$, and if $\mathbf{car} = \{c_1, c_2\}$, $\mathbf{red} = \{c_1\}$, are extensions of the concepts “car” and “red”[4] in the image, then a query requesting “images showing a red car”

[4] For the sake of argument we shall assume that $\mathbf{red}$ be a crisp concept.

would be answered positively on J:

$$\mathbf{some}(\mathbf{car}, \mathbf{red}) \quad = \quad 1\,,$$

while to a query requesting "images in which all cars are red", J would be irrelevant:

$$\mathbf{all}(\mathbf{car}, \mathbf{red}) \quad = \quad 0\,.$$

As another example (taken from the application domain), let us consider a query in which quantification ranges over pixels of image regions. Suppose the user requests for ground temperature maps in which the temperature readings exceed $20°$ Celsius in all of Bavaria. In this case, E is the set of pixel coordinates $E = \{0, \ldots, W - 1\} \times \{0, \ldots, H - 1\}$, where W is the width and H is the height of the temperature map. Suppose the cartographic projection of the map is known. Then, the local region "Bavaria" can be interpreted as the set $B \subseteq E$ of pixel coordinates which belong to Bavaria (relative to the given projection). Further suppose that there is a method which determines the set $T \subseteq E$ of pixel coordinates with an associated temperature reading of more than $20°$ Celsius. Given our definition of $\mathbf{all}$, we can then evaluate $\mathbf{all}(B, T)$. We obtain the result 1 (true) if all pixels which belong to Bavaria also have a $> 20°\,C$ temperature reading (in this case, accept image G for presentation); otherwise, reject G as being irrelevant to the query.

Let us note that – at this stage – we are not yet able to process a query which poses the condition that it is "*hot* in all of Bavaria" because the corresponding concept $\mathbf{hot}$ is fuzzy.

2.2 Fuzzy determiners

Now let us turn to the fuzzy case. Suppose E is some set. A *fuzzy subset* X of E assigns to each $x \in E$ a gradual membership value $\mu_X(x) \in I$, where $I = [0, 1]$. The set of fuzzy subsets of E, denoted by $\widetilde{\mathcal{P}}(E)$, is therefore in one-to-one correspondence with the set of membership functions $\mu : E \longrightarrow I$, i.e. $\widetilde{\mathcal{P}}(E) \cong I^E$. We define fuzzy determiners as follows.

Definition 2 (Fuzzy determiner) *By an n-ary fuzzy determiner on a set E we denote a function $D : \widetilde{\mathcal{P}}(E)^n \longrightarrow I$.*

A fuzzy determiner thus maps any n-tuple $(X_1, \ldots, X_n)$ of fuzzy subsets of E to a corresponding fuzzy membership value $D(X_1, \ldots, X_n) \in I$.

This definition is the obvious generalisation of two-valued determiners (in the sense of TGQ) to the fuzzy case. Nevertheless, these fuzzy determiners pose a problem. We all have strong intuitions about the proper definition of determiners on "crisp" subsets of E; for example, the definition of $\mathbf{all} : \mathcal{P}(E) \times \mathcal{P}(E) \longrightarrow 2$ stated above is certainly uncontroversial. In the fuzzy case, however, such intuitions are lacking: which choice of $\mathbf{all} : \widetilde{\mathcal{P}}(E) \times \widetilde{\mathcal{P}}(E) \longrightarrow I$ should be considered the "correct" one and why?

2.3 Determiner fuzzification schemes

Now, suppose that in order to define a fuzzy determiner $D : \widetilde{\mathcal{P}}(E)^n \longrightarrow I$, we would only have to specify its behaviour on "crisp" arguments, i.e. on $\mathcal{P}(E)^n$, because we had a consistent mechanism $\mathcal{F}$ which generalises our partial definition to a fuzzy determiner $\mathcal{F}(D)$ which is defined on fuzzy arguments also. For example, we could start with $\mathbf{all} : \mathcal{P}(E) \times \mathcal{P}(E) \longrightarrow 2 \subseteq I$ and automatically obtain its fuzzy analogon $\mathcal{F}(\mathbf{all}) : \widetilde{\mathcal{P}}(E) \times \widetilde{\mathcal{P}}(E) \longrightarrow I$. Then the above problem of the proper definition of **all** as a fuzzy determiner would have a unique answer *relative to* $\mathcal{F}$, namely $\mathcal{F}(\mathbf{all})$, and conditions on *reasonable* or *intended* interpretations of fuzzy quantifiers could be stated as axioms governing the behaviour of $\mathcal{F}$. To make this idea work, we firstly introduce fuzzy pre-determiners, and then detail the generalisation mechanism.

Definition 3 (Fuzzy pre-determiner) *An n-ary fuzzy pre-determiner on a set of entities E is a function $D : \mathcal{P}(E)^n \longrightarrow I$.*

Hence in order to define a fuzzy pre-determiner, only results for the application of D to "crisp" arguments (n-tuples of two-valued sets) must be specified. Therefore, an appropriate interpretation of NL determiners is much easier to obtain than in the case of fuzzy determiners. In particular, every two-valued determiner (i.e. determiner in the sense of TGQ, such as **all**), is a fuzzy pre-determiner. Hence we need not guess any new definitions for these determiners.

As examples of fuzzy pre-determiners, let us firstly consider proportional and singular definite determiners:

$$[\geq r\,\%](X_1,\, X_2) = \begin{cases} \frac{1}{2} & : \ \mathrm{card}\, X_1 = 0 \\ 1 & : \ \mathrm{card}\, X_1 \neq 0 \text{ and } \frac{\mathrm{card}\, X_1 \cap X_2}{\mathrm{card}\, X_1} \geq \frac{r}{100} \\ 0 & : \ \text{else} \end{cases}$$

$$\mathbf{the}_{\mathrm{sg}}(X_1,\, X_2) = \begin{cases} \frac{1}{2} & : \ \mathrm{card}\, X_1 \neq 1 \\ 1 & : \ \mathrm{card}\, X_1 = 1 \text{ and } X_1 \subseteq X_2 \\ 0 & : \ \text{else} \end{cases}$$

for all $X_1,\, X_2 \in \mathcal{P}(E)$, $r \in [0,\, 100]$. $\frac{1}{2}$ represents indeterminacy. In the case of $[\geq r\,\%]$ we shall assume that E is finite.

Next let us turn to genuine fuzzy determiners. We shall focus here on three generic examples from which instances such as **many**, **almost all**, **often**, **almost everywhere** etc. can be derived. Let us call these fuzzy pre-determiners $\mathbf{abs\ many}_{\rho,\tau}$ ("many X_1's are X_2 compared to an absolute expected value $\rho \in \mathbb{R}$ with sharpness parameter $\tau \in [0,\, \rho]$"), $\mathbf{rel\ many}_{\rho,\tau}$ ("the proportion of X_1's that are X_2 is large compared to an expected proportion $\rho \in [0,\, 1]$ with sharpness parameter $\tau \in [0,\, \rho]$"), and **as many as possible** (which we interpret as denoting "the proportion of X_1's that are X_2"):

$$\mathbf{abs\ many}_{\rho,\tau}(X_1,\, X_2) = S(\tfrac{\mathrm{card}\, X_1 \cap X_2}{\mathrm{card}\, X_1},\, \rho - \tau,\, \rho + \tau)$$

$$\mathbf{rel\ many}_{\rho,\tau}(X_1,\, X_2) = \begin{cases} \frac{1}{2} & : \ \mathrm{card}\, X_1 = 0 \\ S(\frac{\mathrm{card}\, X_1 \cap X_2}{\mathrm{card}\, X_1},\, \rho - \tau,\, \rho + \tau) & : \ \mathrm{card}\, X_1 \neq 0 \end{cases}$$

$$\mathbf{as\ many\ as\ possible}(X_1,\, X_2) = \begin{cases} \frac{1}{2} & : \ \mathrm{card}\, X_1 = 0 \\ \frac{\mathrm{card}\, X_1 \cap X_2}{\mathrm{card}\, X_1} & : \ \mathrm{card}\, X_1 \neq 0 \end{cases}$$

for all $X_1,\, X_2 \in \mathcal{P}(E)$, where S is Zadeh's S-function [7, pp. 611+].[5]

$\mathbf{abs\ many}_{\rho,\tau}$, $\mathbf{rel\ many}_{\rho,\tau}$ and **as many as possible** cover various meanings of **many** and similar quantifying expressions such as **often**, **relatively often**, **almost all** etc. The parameters $\rho,\, \tau$ are application-specific. Hence as part of a domain model, plausible choices of these parameters must be encoded.

[5] The shape of this curve is as follows: $S(x,\, \rho - \tau,\, \rho + \tau) = 0$ for all $x \in [0,\, \rho - \tau]$. For $x \in [\rho - \tau,\, \rho + \tau]$, S grows monotonically and reaches 1 at $\rho + \tau$. For $x \in [\rho + \tau, 1]$, then, $S(x,\, \rho - \tau,\, \rho + \tau) = 1$. The "indeterminate" value $\frac{1}{2}$ is obtained if x equals the expected value ρ.

Although their global shape should be uncontroversial, there is obviously some degree of freedom in the choice of these fuzzy pre-determiners—some readers might wish to use different definitions. Our theory, however, is in no respect dependent on any particular choice of these fuzzy pre-determiners. By contrast, it is supposed to ensure that *all* fuzzy pre-determiners (including, for example, the alternative definitions of the readers) are systematically generalised to corresponding fuzzy determiners. This is accomplished by means of a determiner fuzzification scheme.

Definition 4 (Determiner fuzzification scheme) *A determiner fuzzification scheme (DFS) $\mathcal{F}$ assigns to each n-ary fuzzy pre-determiner $D : \mathcal{P}(E)^n \longrightarrow I$ an n-ary fuzzy determiner $\mathcal{F}(D) : \widetilde{\mathcal{P}}(E)^n \longrightarrow I$ subject to the axioms DFS 1-10 stated below.*

The DFS axioms are mainly of the "preservation" or "compatibility" (homomorphism) type, i.e. we require that $\mathcal{F}$ preserves relevant properties of a determiner, or that $\mathcal{F}$ is compatible with certain operations for building new determiners from given ones.[6] Hence by stating these axioms, we are making explicit what we think should be expected from a "reasonable" fuzzification scheme. This is of particular importance in the context of fuzzy logic where—even for the propositional connectives—quite a lot of alternative interpretations have been proposed. The DFS-axioms, then, provide a criterion to discern principled approaches to fuzzy quantification from ad-hoc ones.

2.3.1 Correct generalisation

Firstly, we would like that $\mathcal{F}(D)$ coincides with the original fuzzy pre-determiner D when all arguments are crisp subsets of E.

DFS 1 *For every fuzzy pre-determiner $D : \mathcal{P}(E)^n \longrightarrow I$, $\mathcal{F}(D)|_{\mathcal{P}(E)^n} = D$.*

2.3.2 Induced truth functions

The standard choice of negation in fuzzy logic is $\neg x = 1 - x$, for all $x \in I$. With conjunction, there are more common choices (although the standard is certainly $\wedge = \min$). All of these belong to the class of t-norms (cf. [8]), which seems to capture what one would expect of a reasonable conjunction operator. By a canonical construction which we describe now, $\mathcal{F}$ induces a unique fuzzy operator for each of the propositional connectives. We will require that the induced fuzzy truth functions are "reasonable" in the above sense.

Suppose $\{*\}$ is a singleton set. Then $\mathcal{P}(\{*\}) \cong 2 = \{0, 1\}$ by means of the bijection $\kappa : \mathcal{P}(\{*\}) \longrightarrow 2$ defined by $\kappa(\{*\}) = 1$, $\kappa(\varnothing) = 0$. The inverse is $\kappa^{-1}(0) = \varnothing$, $\kappa^{-1}(1) = \{*\}$.

Likewise, $\widetilde{\mathcal{P}}(\{*\}) \cong I$ by means of the bijection $\eta : \widetilde{\mathcal{P}}(\{*\}) \longrightarrow I$, $\eta(X) = \mu_X(*)$ for all $X \in \widetilde{\mathcal{P}}(\{*\})$, with obvious inverse

$$\eta^{-1}(x) \quad = \quad \text{the unique } X \in \widetilde{\mathcal{P}}(\{*\}) \text{ defined by the membership function } \mu_X(*) = x\,,$$

for all $x \in I$.

Now assume $\mathrm{f} : 2^n \longrightarrow 2$ is a propositional function (e.g., $\mathrm{f} = \wedge$). We can view f as a fuzzy pre-determiner $\mathrm{f}^* : \mathcal{P}(\{*\})^n \longrightarrow 2 \subseteq I$ defined by

$$\mathrm{f}^*(X_1, \ldots, X_n) \quad = \quad \mathrm{f}(\kappa(X_1), \ldots, \kappa(X_n))\,.$$

[6] "compatible" means that it does not matter whether we first apply the operation and then apply $\mathcal{F}$, or whether we first apply $\mathcal{F}$ and then apply the considered operation (or its fuzzy analogon).

By applying $\mathcal{F}$, f^* is generalized to a fuzzy determiner $\mathcal{F}(\mathrm{f}^*) : \widetilde{\mathcal{P}}(\{*\})^n \longrightarrow I$, from which we obtain a fuzzy truth function $\widetilde{\mathrm{f}} : I^n \longrightarrow I$ by defining

$$\widetilde{\mathrm{f}}(x_1, \ldots x_n) \quad = \quad \mathcal{F}(\mathrm{f}^*)(\eta^{-1}(x_1), \ldots, \eta^{-1}(x_n))$$

for all $x_1, \ldots, x_n \in I$. Abusing notation, we will again write $\mathcal{F}(\mathrm{f})$ instead of $\widetilde{\mathrm{f}}$, in order to emphasize that $\mathcal{F}$ uniquely determines $\widetilde{\mathrm{f}}$ by the above construction.

DFS 2 *We require:*

a. $\mathcal{F}(\wedge)$ *is a* t*-norm.*

b. $\mathcal{F}(\neg) = \neg$.

We do not impose restrictions on the other propositional connectives because these will result from the other DFS-axioms. When in the following we are using fuzzy connectives, these are always meant to denote the connectives induced by the above construction.

2.3.3 Compatibility with external negation

Assume $D : \mathcal{P}(E)^n \longrightarrow I$ is a fuzzy pre-determiner. $\neg D : \mathcal{P}(E)^n \longrightarrow I$ is defined by

$$(\neg D)(X_1, \ldots, X_n) \quad = \quad \neg(D(X_1, \ldots, X_n)),$$

for all $X_1, \ldots, X_n \in \mathcal{P}(E)$.

In the case of fuzzy determiners, $\neg D$ is defined analogously, choosing $X_1, \ldots, X_n$ from $\widetilde{\mathcal{P}}(E)$.

DFS 3 *For every fuzzy pre-determiner,* $\mathcal{F}(\neg D) = \neg\mathcal{F}(D)$.

2.3.4 Monotonicity

Suppose $D, D' : \mathcal{P}(E)^n \longrightarrow I$ are fuzzy pre-determiners of the same arity. Let us write $D \leq D'$ iff for all $X_1, \ldots, X_n \in \mathcal{P}(E)$,

$$D(X_1, \ldots, X_n) \quad \leq \quad D'(X_1, \ldots, X_n).$$

(On fuzzy determiners, define $\leq$ analogously).

We require $\mathcal{F}$ to be monotonic, i.e. to preserve inequations between fuzzy pre-determiners:

DFS 4 *For all fuzzy pre-determiners* $D, D' : \mathcal{P}(E)^n \longrightarrow I$*, if* $D \leq D'$*, then also* $\mathcal{F}(D) \leq \mathcal{F}(D')$.

Note. The converse implication "$\Leftarrow$" follows from DFS-1.

2.3.5 Compatibility with argument transposition

Assume A is some set and $n \in \mathbb{N}$. By $\tau_i : A^n \longrightarrow A^n$ $(1 \leq i \leq n)$ we denote the function defined by

$$\tau_i(a_1, \ldots, a_n) \quad = \quad (a_1, \ldots, a_{i-1}, a_n, a_{i+i}, \ldots, a_{n-1}, a_i),$$

which swaps the i-th and n-th component in $(a_1, \ldots, a_n) \in A^n$. By $\circ$ we denote function composition.

DFS 5 *For every n-ary fuzzy pre-determiner D and all $1 \leq i \leq n$, $\mathcal{F}(D \circ \tau_i) = \mathcal{F}(D) \circ \tau_i$.*

Note. By a finite number of such argument transpositions, every permutation of arguments of D can be obtained. Hence this axiom states that $\mathcal{F}$ treats all arguments of determiners in essentially the same way.

2.3.6 Compatibility with internal complementation

For every fuzzy pre-determiner $D : \mathcal{P}(E)^n \longrightarrow I$, we define $D\neg : \mathcal{P}(E)^n \longrightarrow I$ by

$$D\neg(X_1, \ldots, X_n) \quad = \quad D(X_1, \ldots, X_{n-1}, \neg X_n)$$

for all $X_1, \ldots, X_n \in \mathcal{P}(E)$, where $\neg X_n$ denotes complementation. Again, we use an analog definition based on $\widetilde{\mathcal{P}}(E)$ in the case of fuzzy determiners. (For a fuzzy subset $X_n \in \widetilde{\mathcal{P}}(E)$, the complement $\neg X_n$ is defined by $\mu_{\neg X_n}(x) = 1 - \mu_{X_n}(x)$ for all $x \in E$).

DFS 6 *For every fuzzy pre-determiner, $\mathcal{F}(D\neg) = \mathcal{F}(D)\neg$.*

Notes

- Due to the axiom of argument transposition, this generalizes to complementation in arbitrary argument positions.
- By de Morgan's law, $x \vee y = \neg(\neg x \wedge \neg y)$. By applying DFS-2, DFS-3 and DFS-6, we obtain that $\mathcal{F}(\vee)$ is the complementary co-t-norm of $\mathcal{F}(\wedge)$.

2.3.7 Compatibility with internal meets

Now assume $D : \mathcal{P}(E)^n \longrightarrow I$ is a fuzzy pre-determiner. By $D\cap : \mathcal{P}(E)^{n+1} \longrightarrow I$ we denote the fuzzy pre-determiner defined by

$$D\cap(X_1, \ldots X_{n+1}) \quad = \quad D(X_1, \ldots, X_{n-1}, X_n \cap X_{n+1})$$

for all $X_1, \ldots, X_{n+1} \in \mathcal{P}(E)$. (Similar definition for fuzzy determiners. In this case, the meet of $X_n, X_{n+1} \in \widetilde{\mathcal{P}}(E)$ is defined by $\mu_{X_n \cap X_{n+1}}(x) = \mu_{X_n}(x) \wedge \mu_{X_{n+1}}(x)$ for all $x \in E$, where $\wedge$ is the t-norm induced by $\mathcal{F}$.)

DFS 7 *For all fuzzy pre-determiners, $\mathcal{F}(D\cap) = \mathcal{F}(D)\cap$.*

Notes

- By iterating and transposing arguments, this axiom generalises to arbitrary meets of non-redundant tuples of arguments.
- From the axiom of internal complementation, we obtain $\mathcal{F}(D\cup) = \mathcal{F}(D)\cup$, where the join operation is defined analogously.

2.3.8 Compatibility with insertion of constants

For every fuzzy pre-determiner $D : \mathcal{P}(E)^{n+1} \longrightarrow I$ and every $A \in \mathcal{P}(E)$, we denote by $D{\downarrow}A$ the n-ary fuzzy pre-determiner defined by

$$(D{\downarrow}A)(X_1, \ldots, X_n) \quad = \quad D(X_1, \ldots, X_n, A),$$

for all $X_1, \ldots, X_n \in \mathcal{P}(E)$ (analogously for fuzzy determiners).

DFS 8 *For every fuzzy pre-determiner $D : \mathcal{P}(E)^{n+1} \longrightarrow I$ and every $A \in \mathcal{P}(E)$, $\mathcal{F}(D{\downarrow}A) = \mathcal{F}(D){\downarrow}A$.*

2.3.9 Preservation of monotonicity

We also require that $\mathcal{F}$ preserves monotonicity properties of a determiner. To be precise, assume $D : \mathcal{P}(E)^n \longrightarrow I$ is a fuzzy pre-determiner. We say that D is monotonically increasing in its i-th argument iff for all $X_1, \ldots, X_n, X_i' \in \mathcal{P}(E)$ with $X_i \subseteq X_i'$,

$$D(X_1, \ldots, X_n) \leq D(X_1, \ldots, X_{i-1}, X_i', X_{i+1}, X_n).$$

For the definition on n-ary fuzzy determiners, we choose fuzzy subsets of $\widetilde{\mathcal{P}}(E)$.

DFS 9 *For every n-ary fuzzy per-determiner D, if D is monotonically increasing in its n-th argument, then so is $\mathcal{F}(D)$.*

Notes

- By argument transpositions, $\mathcal{F}$ preserves (increasing) monotonicity in all arguments of a determiner.
- By DFS-9 and the axiom of external negation DFS-3, $\mathcal{F}$ also preserves decreasing monotonicity.

2.3.10 Compatibility with functional application

Let E, E' be any sets. Every function $\mathrm{f} : E \longrightarrow E'$ induces a function $\mathcal{P}(E) \longrightarrow \mathcal{P}(E')$ (again denoted by f) which maps subsets X of E onto their corresponding images $\mathrm{f}(X) = \{\mathrm{f}(x) : x \in X\}$. In the fuzzy case, $\mathrm{f} : \widetilde{\mathcal{P}}(E) \longrightarrow \widetilde{\mathcal{P}}(E')$ is obtained from Zadeh's *extension principle* [9], which states that for all $X \in \widetilde{\mathcal{P}}(E)$, $\mathrm{f}(X)$ is the fuzzy subset of E' defined by the membership function

$$\mu_{\mathrm{f}(X)}(y) \quad = \quad \sup\{\mu_X(x) : x \in E \text{ and } \mathrm{f}(x) = y\}.$$

DFS 10 *Assume $D : \mathcal{P}(E')^n \longrightarrow I$ is a fuzzy pre-determiner and $\mathrm{f}_1, \ldots, \mathrm{f}_n : E \longrightarrow E'$ are functions. Define $D' : \mathcal{P}(E)^n \longrightarrow I$ by*

$$D'(X_1, \ldots, X_n) \quad = \quad D(\mathrm{f}_1(X_1), \ldots, \mathrm{f}_n(X_n)),$$

for all $X_1, \ldots, X_n \in \mathcal{P}(E)$. Then for all $X_1, \ldots, X_n \in \widetilde{\mathcal{P}}(E)$,

$$\mathcal{F}(D')(X_1, \ldots, X_n) \quad = \quad \mathcal{F}(D)(\mathrm{f}_1(X_1), \ldots, \mathrm{f}_n(X_n)).$$

Notes

- Basically, the axiom of functional application states that it does not matter whether we first generalise (using $\mathcal{F}$) and then apply the f's, or vice versa.
- One is tempted to require a generalisation to m-ary functions $\mathrm{f}_i : E^{m_i} \longrightarrow E'$ where $m_i > 1$, i.e. to
$$D'(\mathrm{f}_1(X_{j_1^1}, \ldots, X_{j_{m_1}^1}), \ldots, \mathrm{f}_n(X_{j_1^n}, \ldots, X_{j_{m_n}^n}))\,.$$
However, this would imply the law of "tertium non datur" and hence exclude all interesting t-norms (in particular, $\min$) from DFS-2a.
- Assume $\pi : E \longrightarrow E'$ is a bijection and
$$D'(X_1, \ldots, X_n) \quad = \quad D(\pi(X_1), \ldots, \pi(X_n))\,.$$
Then we obtain
$$\mathcal{F}(D')(X_1, \ldots, X_n) \quad = \quad \mathcal{F}(D)(\pi(X_1), \ldots, \pi(X_n))\,.$$
From this we see that $\mathcal{F}$ may not depend on any particular properties of elements of E.

3 A model of the DFS axioms

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 practical purposes), we will now present an actual model.

Assume $X \in \widetilde{\mathcal{P}}\,(E)$ is a fuzzy subset of some set E and $\alpha \in I$. By $(X)_{\geq\alpha} \in \mathcal{P}\,(E)$ we denote the α-cut
$$(X)_{\geq\alpha} \quad = \quad \{X \in E : \mu_X(x) \geq \alpha\}\,,$$
by $(X)_{>\alpha} \in \mathcal{P}\,(E)$ the strict α-cut
$$(X)_{>\alpha} \quad = \quad \{x \in E : \mu_X(x) > \alpha\}\,.$$
As a first attempt to find a DFS, let us consider
$$\mathcal{A}(D)(X_1, \ldots, X_n) \quad = \quad \int_0^1 D((X_1)_{\geq\alpha}, \ldots, (X_n)_{\geq\alpha})\,\mathrm{d}\alpha\,,$$
where $D : \mathcal{P}\,(E)^n \longrightarrow I$ is a fuzzy pre-determiner and $X_1, \ldots, X_n \in \widetilde{\mathcal{P}}\,(E)$.

$\mathcal{A}$ fails to be a DFS as the above integral may not exist (we have not imposed any restrictions on the "well-behavedness" of D).[7] Even in the case that $\mathcal{A}$ is defined, it fails to satisfy the axiom of internal complementation (DFS-6) because $(\neg X)_{\geq\alpha} = \neg(X)_{>1-\alpha}$, which is different from $\neg(X_{\geq\alpha})$ in most cases ($\neg$ denotes complementation).

[7]To see this, choose a non-measurable function $\mathrm{f} : I \longrightarrow I$ and let $E = I$. Define $D : \mathcal{P}\,(I) \longrightarrow I$ by $D(X) = \mathrm{f}(\inf X)$ for all $X \in \mathcal{P}\,(I)$. $\mathcal{A}(D)$ is undefined on the fuzzy subset $X^* \in \widetilde{\mathcal{P}}\,(I)$ defined by $\mu_{X^*}(x) = x$ for all $x \in I$, because $D((X^*)_{\geq\alpha}) = \mathrm{f}(\alpha)$.

The DFS we are actually using in our system is based on the idea of using “three-valued” cuts (as opposed to the two-valued α-cuts).

We therefore reconsider Kleene's three-valued logic (see e.g. [10, p. 29]). Kleene's logic can be obtained from two-valued logic by the following mechanism of generalizing a propositional function $\mathrm{f} : 2^n \longrightarrow 2$ to the three-valued $\widehat{\mathrm{f}} : \{0, \frac{1}{2}, 1\}^n \longrightarrow \{0, \frac{1}{2}, 1\}$.

Suppose $x_1, \ldots, x_n \in \{0, \frac{1}{2}, 1\}$ are given. Associate to each x_i a set $Y_i \subseteq \mathcal{P}(2)$ as follows:

$$Y_i = \begin{cases} \{0\} & : \quad x_i = 0 \\ \{0, 1\} & : \quad x_i = \frac{1}{2} \\ \{1\} & : \quad x_i = 1 \end{cases}$$

So the “indeterminate” value $\frac{1}{2}$ is treated by considering both alternatives $0, 1$. The truth-values 0 and 1 do not induce any indeterminacy.

Then define $\widehat{\mathrm{f}} : \{0, \frac{1}{2}, 1\}^n \longrightarrow \{0, \frac{1}{2}, 1\}$ by

$$\widehat{\mathrm{f}}(x_1, \ldots, x_n) = \begin{cases} 1 & : \quad \mathrm{f}(y_1, \ldots, y_n) = 1 \text{ for all } y_i \in Y_i,\ i = 1, \ldots, n \\ 0 & : \quad \mathrm{f}(y_1, \ldots, y_n) = 0 \text{ for all } y_i \in Y_i,\ i = 1, \ldots, n \\ \frac{1}{2} & : \quad \text{else} \end{cases}$$

for all $x_1, \ldots, x_n \in \{0, \frac{1}{2}, 1\}$.

We recall the definition of fuzzy median,

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

$\mathrm{m}_{\frac{1}{2}}$ is an associative mean operator [11] and the only stable associative symmetric sum [12]. Due to its associativity and commutativity, $\mathrm{m}_{\frac{1}{2}}$ can be generalized to arbitrary finite sets of arguments. Noting that for all finite $X = \{x_1, \ldots, x_n\} \subseteq I$, $n \geq 2$, it holds that

$$\mathrm{m}_{\frac{1}{2}} X = \mathrm{m}_{\frac{1}{2}} (\min X, \max X) ,$$

the proper definition of $\mathrm{m}_{\frac{1}{2}} X$ in the case $n = 0, n = 1$ is

$$\mathrm{m}_{\frac{1}{2}} \varnothing = \mathrm{m}_{\frac{1}{2}}(\min \varnothing, \max \varnothing) = \mathrm{m}_{\frac{1}{2}}(1, 0) = \tfrac{1}{2} ,$$

$$\mathrm{m}_{\frac{1}{2}}\{u\} = \mathrm{m}_{\frac{1}{2}}(\min\{u\}, \max\{u\}) = \mathrm{m}_{\frac{1}{2}}(u, u) = u .$$

We extend $\mathrm{m}_{\frac{1}{2}}$ to arbitrary subsets $X \subseteq I$ by defining

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

for all $X \subseteq I$, which is obviously compatible with the definition on finite subsets of I.

By using the fuzzy median $\mathrm{m}_{\frac{1}{2}}$, we can now state the above definition more compactly,

$$\widehat{\mathrm{f}}(x_1, \ldots, x_n) = \mathrm{m}_{\frac{1}{2}}\{\mathrm{f}(y_1, \ldots, y_n) : y_i \in Y_i,\ i = 1, \ldots, n\} .$$

This median-based definition has the advantage that it can also be applied to functions $\mathrm{f} : 2^n \longrightarrow I$, yielding $\widehat{\mathrm{f}} : \{0, \frac{1}{2}, 1\}^n \longrightarrow I$.

Using the same mechanism, we can generalize two-valued determiners $D : \mathcal{P}(E)^n \longrightarrow 2$ to three-valued determiners $\widehat{D} : \widehat{\mathcal{P}}(E)^n \longrightarrow \{0, \frac{1}{2}, 1\}$, where $\widehat{\mathcal{P}}(E) \cong \{0, \frac{1}{2}, 1\}^E$ is the set of all three-valued subsets of E (each three-valued subset X of E is uniquely determined by its membership function $\nu_X : E \longrightarrow \{0, \frac{1}{2}, 1\}$).

Suppose $D : \mathcal{P}(E)^n \longrightarrow 2$ and $X_1, \ldots, X_n \in \widehat{\mathcal{P}}(E)$ are given. To each of the three-valued sets $X_i \in \widehat{\mathcal{P}}(E)$ we associate a set of two-valued sets $\mathcal{Y}_i \subseteq \mathcal{P}(E)$,

$$\mathcal{Y}_i = \{Y \subseteq E : X_i^{\min} \subseteq Y \subseteq X_i^{\max}\}$$

where

$$\begin{aligned} X_i^{\min} &= \{x \in E : \nu_{X_i}(x) = 1\}, \\ X_i^{\max} &= \{x \in E : \nu_{X_i}(x) = \tfrac{1}{2} \text{ or } \nu_{X_i}(x) = 1\} \end{aligned}$$

So if $x \in E$ such that $\mu_{X_i}(x) = 1$, then all $Y \in \mathcal{Y}_i$ contain x; if $x \in E$ such that $\mu_{X_i}(x) = 0$, then no $Y \in \mathcal{Y}_i$ contains x, and for those $x \in E$ with $\mu_{X_i} = \frac{1}{2}$, $\mathcal{Y}_i$ contains all combinations of the alternatives $x \in Y$, $x \notin Y$.

Using $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, we can now define

$$\widehat{D}(X_1, \ldots, X_n) = \begin{cases} 1 & : D(Y_1, \ldots, Y_n) = 1 \text{ for all } Y_i \in \mathcal{Y}_i,\ i = 1, \ldots, n \\ 0 & : D(Y_1, \ldots, Y_n) = 0 \text{ for all } Y_i \in \mathcal{Y}_i,\ i = 1, \ldots, n \\ \frac{1}{2} & : \text{else} \end{cases}$$

Again, we can restate this

$$\widehat{D}(X_1, \ldots, X_n) = \mathrm{m}_{\frac{1}{2}}\{D(Y_1, \ldots, Y_n) : Y_i \in \mathcal{Y}_i,\ i = 1, \ldots, n\}.$$

The $\mathrm{m}_{\frac{1}{2}}$-based definition can also be used to generalize from $D : \mathcal{P}(E)^n \longrightarrow I$ to $\widehat{D} : \widehat{\mathcal{P}}(E)^n \longrightarrow I$.

Now let us introduce three-valued cuts on fuzzy sets in order to build on these definitions. Assume $X \in \widetilde{\mathcal{P}}(E)$ is a fuzzy subset of some set E and $\gamma \in (0, 1]$. Denote the membership function of X by $\mu_X : E \longrightarrow I$. We can cut X to a three-valued set $\widehat{X} \in \widehat{\mathcal{P}}(E)$ with membership function $\nu_{\widehat{X}} = \mu_X^\gamma : E \longrightarrow \{0, \frac{1}{2}, 1\}$ by setting

$$\mu_X^\gamma(x) = \begin{cases} 0 & : \mu_X(x) \le \frac{1}{2} - \frac{1}{2}\gamma \\ \frac{1}{2} & : \frac{1}{2} - \frac{1}{2}\gamma < \mu_X(x) < \frac{1}{2} + \frac{1}{2}\gamma \\ 1 & : \mu_X(x) \ge \frac{1}{2} + \frac{1}{2}\gamma \end{cases}$$

for all $x \in E$. For $\gamma = 0$, we define

$$\mu_X^0(x) = \begin{cases} 0 & : \mu_X(x) < \frac{1}{2} \\ \frac{1}{2} & : \mu_X(x) = \frac{1}{2} \\ 1 & : \mu_X(x) > \frac{1}{2} \end{cases}$$

for all $x \in E$.

γ can be thought of as a parameter of "cautiousness". If $\gamma = 0$, the set of indeterminates contains only those $x \in E$ with $\mu_X(x) = \frac{1}{2}$; 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 $\mu_X(x) \in \{0, 1\}$ are interpreted as indeterminates.

Now let $D : \mathcal{P}(E)^n \longrightarrow I$ a fuzzy pre-determiner and $\gamma \in I$. Define $D_\gamma : \widetilde{\mathcal{P}}(E) \longrightarrow I$ by

$$D_\gamma(X_1, \ldots, X_n) = \mathrm{m}_{\frac{1}{2}}\{D(Y_1, \ldots, Y_n) : Y_i \in \mathcal{Y}_i^{\gamma}, i = 1, \ldots, n\},$$

where $\mathcal{Y}_i^\gamma$ is obtained from the three-valued cut of X_i under γ ($i = 1, \ldots, n$), i.e.

$$\begin{aligned}
\mathcal{Y}_i^\gamma &= \{Y \subseteq E : (X_i)_\gamma^{\min} \subseteq Y \subseteq (X_i)_\gamma^{\max}\} \\
(X_i)_\gamma^{\min} &= \begin{cases} (X_i)_{\geq \frac{1}{2} + \frac{1}{2}\gamma} & : \gamma \in (0, 1] \\ (X_i)_{> \frac{1}{2}} & : \gamma = 0 \end{cases} \\
(X_i)_\gamma^{\max} &= \begin{cases} (X_i)_{> \frac{1}{2} - \frac{1}{2}\gamma} & : \gamma \in (0, 1] \\ (X_i)_{\geq \frac{1}{2}} & : \gamma = 0 \end{cases}
\end{aligned}$$

This assignment $D : \mathcal{P}(E)^n \longrightarrow I \;\mapsto\; D_\gamma : \widetilde{\mathcal{P}}(E)^n \longrightarrow I$ is not a DFS yet; the fuzzy median suppresses to much structure. Notably, $\neg_\gamma \neq \neg$ for all $\gamma \in I$, i.e. DFS-2b does not hold. We must hence take into account the results obtained at each level of cautiousness, which we accomplish by integrating over γ:

$$\mathcal{M}(D)(X_1, \ldots, X_n) = \int_0^1 D_\gamma(X_1, \ldots, X_n)\,\mathrm{d}\gamma\,.$$

Firstly, let us note that the integral is well-defined, regardless of D. Provided a choice of $X_1, \ldots, X_n$ in $\widetilde{\mathcal{P}}(E)$, integration ranges over the function $D(\_) : I \longrightarrow I$ defined by

$$D(\gamma) = D_\gamma(X_1, \ldots, X_n)$$

for all $\gamma \in I$. Now if $D(0) \geq \frac{1}{2}$, then $D(\_)$ is monotonically decreasing in γ; if $D(0) \leq \frac{1}{2}$, then $D(\_)$ is monotonically increasing in γ. Hence, as every monotonic function on a closed interval, $D(\_)$ is Riemann integrable.

$\mathcal{M}$ has some nice theoretical properties. In particular, $\mathcal{M}$ is a DFS.[8]

Furthermore, $\mathcal{M}(\wedge) = \min$, $\mathcal{M}(\vee) = \max$ and $\mathcal{M}(\neg) = 1 - x$. Hence by integrating over "levels of cautiousness", we have captured the essence of fuzzy logic in a single formula!

For the well-known quantors defined on page 4, the following results are obtained (E finite):

$$\begin{aligned}
\mathcal{M}(\forall)(X) &= \min_{x \in E} \mu_X(x) \\
\mathcal{M}(\exists)(X) &= \max_{x \in E} \mu_X(x) \\
\mathcal{M}(\mathbf{some})(X_1, X_2) &= \max_{x \in E} \min(\mu_{X_1}(x), \mu_{X_2}(x)) \\
\mathcal{M}(\mathbf{all})(X_1, X_2) &= \min_{x \in E} \max(1 - \mu_{X_1}(x), \mu_{X_2}(x)) \\
\mathcal{M}(\mathbf{atleast\ n})(X_1, X_2) &= \mu_{(n)}\,,
\end{aligned}$$

where $\mu_{(n)}$ is the n-th largest element in the ordered sequence of membership values of $X_1 \cap X_2$ (including duplicates).

[8] Proof: Glöckner 1997 [13].

In the general case, $\mathcal{M}(D)$ cannot be described by such closed-form expressions. However, $\mathcal{M}(D)$ always has an efficient operationalisation on the basis of histogram computations if D is quantitative.[9] In the experimental HPQS system, we have implemented 28 fuzzy determiners in this fashion. Implementation aspects are described in [14].

4 Discussion of DFS

This axiomatic account of fuzzy quantification offers clear advantages compared to the approaches described in the literature. Ever since Zadeh published his basic papers on fuzzy quantification [15, 7, 16], fuzzy quantifiers have generally been treated as fuzzy subsets $Q \in \widetilde{\mathcal{P}}(\mathbb{R}^+)$ of the non-negative reals (so-called "absolute quantifiers") or of the unit interval ($Q \in \widetilde{\mathcal{P}}(I)$, "proportional quantifiers"). In order to make these fuzzy numbers applicable to fuzzy sets for the purpose of quantification, a mechanism (which we denote by $\mathcal{C}$) is needed which transports Q to a fuzzy determiner

$$\mathcal{C}(Q) : \widetilde{\mathcal{P}}(E) \longrightarrow I \quad \text{(unrestricted use, relative to } E\text{), or}$$
$$\mathcal{C}(Q) : \widetilde{\mathcal{P}}(E) \times \widetilde{\mathcal{P}}(E) \longrightarrow I \quad \text{(restricted use, relative to first argument).}$$

Zadeh has also formulated the idea that in order to evaluate a statement "$Q\, X's$ are A" (in our notation: to compute $\mathcal{C}(Q)(A)$), one should instead evaluate the statement "$\mathrm{card}\, X$ is Q", where card is a scalar or fuzzy measure of the cardinality of the fuzzy set $A \in \widetilde{\mathcal{P}}(E)$ associated with the linguistic variable X. The approaches described in the literature mainly differ in the measure of fuzzy cardinality used and in the way that the required comparison of fuzzy cardinalities is accomplished. Zadehs original Σ-count approach was shown to be implausible (see e.g. Ralescu [17, 18]). Ralescu's possibilistic approach and Yagers 1984 [19] as well as his later OWA-approach [20, 21, 22], however, have similar deficiencies (notably, neither provides a convincing account of the "restricted quantification" case stated above).

Firstly, these approaches are based on the assumption that the base set E be finite. DFS, however, requires a proper treatment of quantification over infinite base sets as well (and our model $\mathcal{M}$ can handle this case).

Next, although in all these approaches, the term "linguistic quantifier" is used, none of them is compatible with the results of modern linguistic theory. For example, the Theory of Generalized Quantifiers wisely discerns determiners of different arities n because there are genuine multiplace determiners such as *more* in "more men than women are smokers" ($n = 3$). These cannot be interpreted as absolute or proportional quantifiers as they are known to be irreducible to two-place determiners (Hamm [23, pp. 23+]).

Although absolute and proportional determiners form natural subclasses, it should be noted that not every determiner (in the sense of TGQ) corresponds to these classes, not even every two-place determiner. Quite the reverse, the comprehensive classification of natural language determiners by Keenan and Stavi [6, pp. 253-256] distinguishes 16 main classes of determiners (among others, determiners of exception like $\mathbf{all\ except\ one}$, definites like $\mathbf{the\ ten}$, bounding determiners $\mathbf{only}$,...). It should be noted that TGQ can model all these very different classes of determiners with the same basic mechanism, just as our DFS approach does: by a single DFS, *every* determiner in the sense of TGQ can be generalized to a fuzzy determiner which also accepts fuzzy subsets of the universe E as its arguments.

[9] i.e. if D depends only on the cardinalities of boolean combinations of its arguments.

By contrast, the fuzzy approaches to quantification cited above rely on the distinction of proportional and absolute determiners. There is no simple or straightforward way to generalise these approaches to arbitrary determiners; just as there is no uniform account of cardinal (absolute) and proportional determiners in these approaches.

DFS, however, is a compatible generalisation of contemporary linguistic theory. Its axiomatic foundation ensures that the generalisations produced by a given DFS are intuitively plausible. The uniform approach (one DFS is sufficient) also guarantees the desired systematicity.

Another advantage is that by choosing a DFS, a unique definition of all fuzzy connectives is induced. For example, each DFS is compatible with exactly one choice of conjunction, disjunction, implication, weighted conjunction and weighted disjunction—there is no degree of freedom left in the definition of these connectives.

5 Applications

5.1 Evaluation of quantified expressions in NL queries

Our theory has been designed to establish a principled account of quantified expressions in NL queries, and we have ensured this by stating it in the form of axioms.

As we have also provided an actual model $\mathcal{M}$ of the theory, we are now able to evaluate arbitrary fuzzy quantifying expressions on fuzzy subsets of arbitrary base sets. We only need to provide a fuzzy pre-determiner D which reflects the semantics of the considered expression on "crisp" sets, and then implement the corresponding fuzzy determiner $\mathcal{M}(D)$.

In the HPQS system, the fuzzy quantifiers $\mathcal{M}(D)$ thus obtained are used to evaluate NL queries involving quantified expressions which range over sets of time points, pixel coordinates, and image regions.

We will now reconsider the problem of searching for ground temperature maps subject to the condition that "it is hot in all of Bavaria".

Again suppose $G : E \longrightarrow \mathbb{R}$ is the ground temperature image under consideration, where E is the set of pixel coordinates, and $B \in \mathcal{P}(E) \subseteq \widetilde{\mathcal{P}}(E)$ is the (crisp) extension of Bavaria in E. Further assume that the fuzzy concept "hot" is modelled by some fuzzy subset $\mathbf{hot}$ in $\widetilde{\mathcal{P}}(\mathbb{R})$. Then from G, we obtain the fuzzy region H of pixels which are classified as "hot" by

$$\mu_H(x) = \mu_{\mathbf{hot}}(G(x))$$

for all $x \in E$. Now as B is crisp, evaluation of $\mathcal{M}(\mathbf{all})$ yields

$$\mathcal{M}(\mathbf{all})(B,\, H) \quad = \quad \min_{x \in B} \mu_H(x)\,,$$

i.e. the minimal degree of "hotness" which a pixel belonging to Bavaria obtains in the temperature map G. From this we see that—when applied to image regions—the classical quantifiers $\mathbf{all}$ and $\mathbf{some}$ are very sensitive to noise in the pixel values (a single erroneous pixel might change the result drastically). $\mathcal{M}$ cannot be blamed for this because this "brittleness" is already present in the two-valued determiners $\mathbf{all}$ and $\mathbf{some} : \mathcal{P}(E) \times \mathcal{P}(E) \longrightarrow \mathbf{2}$ from which we have started. Genuine fuzzy quantifiers, e.g. of the $\mathbf{abs\ many}$ or $\mathbf{rel\ many}$ type, are much more robust in this respect.

5.2 Quantifiers as aggregation operators in textual information retrieval

Since in colloquial language, people are using natural language quantifiers with the same ease as the connectives "and", "or", and "not", these quantifiers constitute a rich yet *human understandable* class of operators for information aggregation and data fusion. For example, they can be used in textual information retrieval—in a similar fashion as the operators "and" and "or" in traditional query languages—to provide for more natural (e.g., compromising) ways of aggregating over sets of search terms.

Suppose T is a set of search terms and $B,\, R_d \in \widetilde{\mathcal{P}}(T)$ are fuzzy subsets of T. Assume $\mu_B(t)$ describes the relevance of term $t \in T$ to the user's information need[10] and $\mu_{R_d}(t)$ describes the relevance of term t with respect to a document d under consideration (such a gradual term-document relevance could e.g. arise as the result of an approximate match). Now let $Q : \mathcal{P}(T) \times \mathcal{P}(T) \longrightarrow I$ be a two-place fuzzy pre-determiner. By computing $\mathcal{M}(Q)(B,\, R_d)$, Q can be used to aggregate the term-document relevances $\mu_{R_d}(t)$, $t \in T$ relative to the user's interest $\mu_B(t)$. The choice of $Q \in \{\mathbf{atleast\ n},\, \mathbf{more\ than\ r}\,\%,\, \mathbf{many},\, \mathbf{few},\, \mathbf{as\ many\ as\ possible},\, \ldots\}$ depends on the desired specificity of search and the desired mode of aggregation.

5.3 Focussing of thesaurus search

Fuzzy quantifiers can also prove useful when applied to thesaurus search, as e.g. in

> *Show me all documents which contain as many terms associated with "information retrieval" as possible!*

In this case, $B(t) = \mu_{\mathrm{Th}}(t,\, t')$ where $\mathrm{Th} \in \widetilde{\mathcal{P}}(T \times T)$ is a fuzzy thesaurus [24, 25], i.e. a reflexive and symmetric fuzzy relation judging association strength between pairs of search terms, and $t' = \mathbf{information\ retrieval}$. This kind of "thesaurus search with quantifiers" yields more specific and purposive results than the usual "thesaurus search by disjunction" of associated search terms.

5.4 Fitting the concept of fuzzy quantifiers to the HPQS

In order to show how these quantifiers may be used directly in the evaluation methods for semantic search, we consider the following sample query, which can be posed to our system:

> *Show me pictures of cloud formation over Bavaria taken in August 1997!*

This query requests weather images C from the database with an associated date $C.\mathrm{date}$ in August 1997 subject to the condition $R(C)$

> C *witnesses a weather situation of cloud formation over Bavaria.*

So while the user simply requests *documents* of the database (as opposed to question-answering in question-answer systems and expert systems), the condition R imposed on these documents refers to an underlying *domain model* (in this case, of meteorology) in which the concept of "cloud formation" over some specified region is given a procedural interpretation.

[10] The user might provide these weights by annotating search terms

Considering a single image C is not sufficient to compute $R(C)$ because “formation” is essentially a matter of change and thus can only be judged relative to other images (in this case, images of relevance to the weather situation *before* $C.\mathrm{date}$). Therefore, at least the weather image immediately preceeding C must be taken into account.

The task of detecting the relevant documents is decomposed by the retrieval component as follows: “cloud formation” in a given local region is interpreted as a strong increase in cloudiness in that region in two consecutive weather maps. The systems thus scans through the sequence of images in the specified time interval (August 1997). For each image C under consideration, it firstly determines the predecessor image P. The following operations are then applied to C and P :

a) compute estimates $C_1,\ P_1$ of the density of low clouds in C and P;

b) transform the results into the cartographic projection of the region B under consideration – in this case, $B = \mathbf{Bavaria}$, yielding C_2, P_2;

c) compute the fuzzy evaluations (grey-value images) $C_3 = \mathbf{cloudy}(C_2)$, $P_3 = \neg\mathbf{cloudy}(P_2)$

d) combine C_3 and P_3 *relative to the given region* B by a suitable quantifier to form the gradual result R:

$$R \quad = \quad \min\{\,\mathcal{M}(\geq 70\,\%)(B,\,C_3),\ \mathcal{M}(\geq 70\,\%)(B,\,P_3)\,\}\ .$$

Hence R is the fuzzy conjunction of the conditions that more than $70\,\%$ of Bavaria are sunny in the predecessor image P and that more than $70\,\%$ of Bavaria are cloudy in the current image C.[11]

Intermediate results of the search process are shown in fig. 2. As indicated by the search results in fig. 3, the system detects images in which there is a clearly-visible increase in cloudiness.

[11]The choice of these percentages is of course to some degree arbitrary. However it should be noted that each reasonable choice yields an improvement compared to an approach which does not utilize such modelling tools.

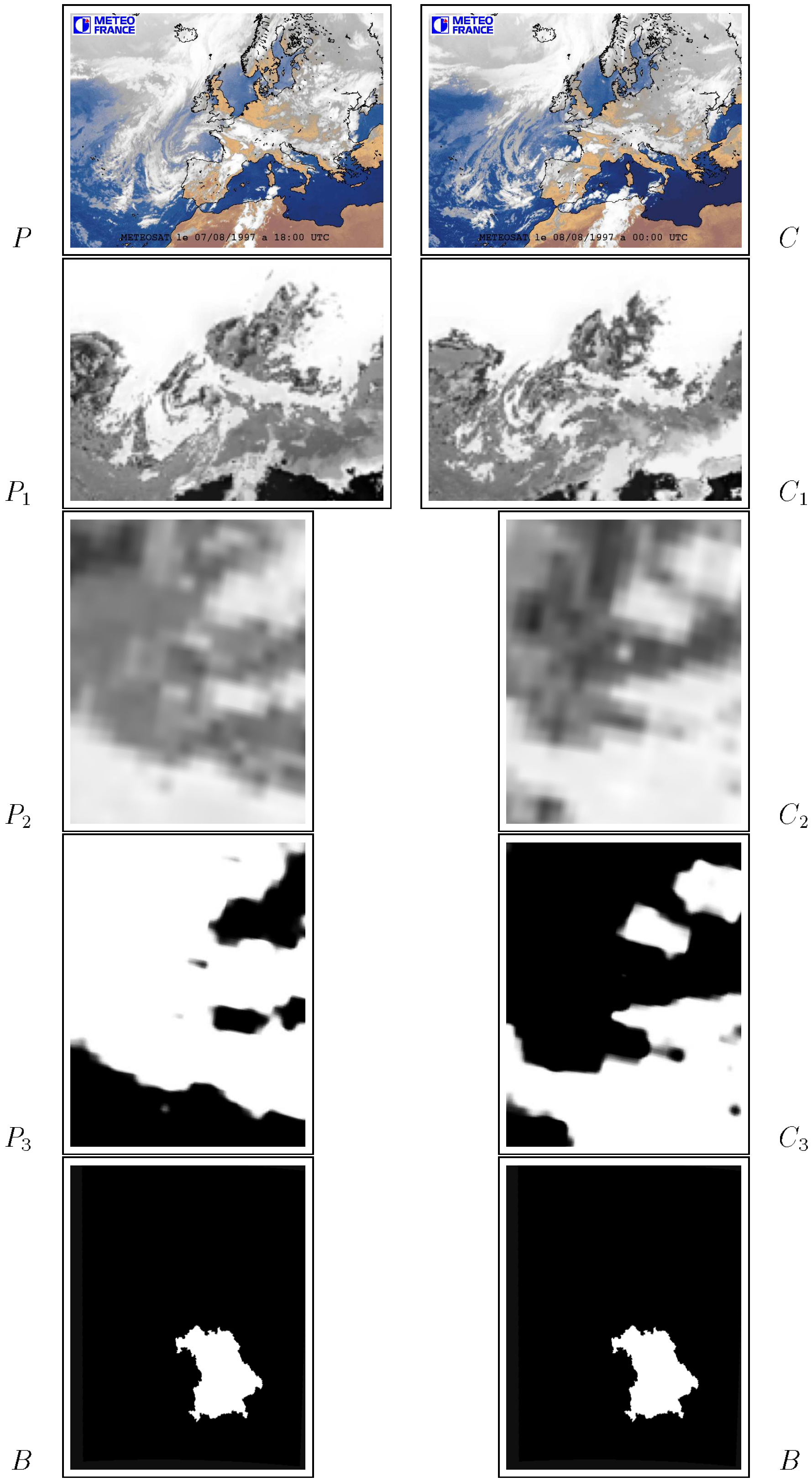

Figure 2: Intermediate search results

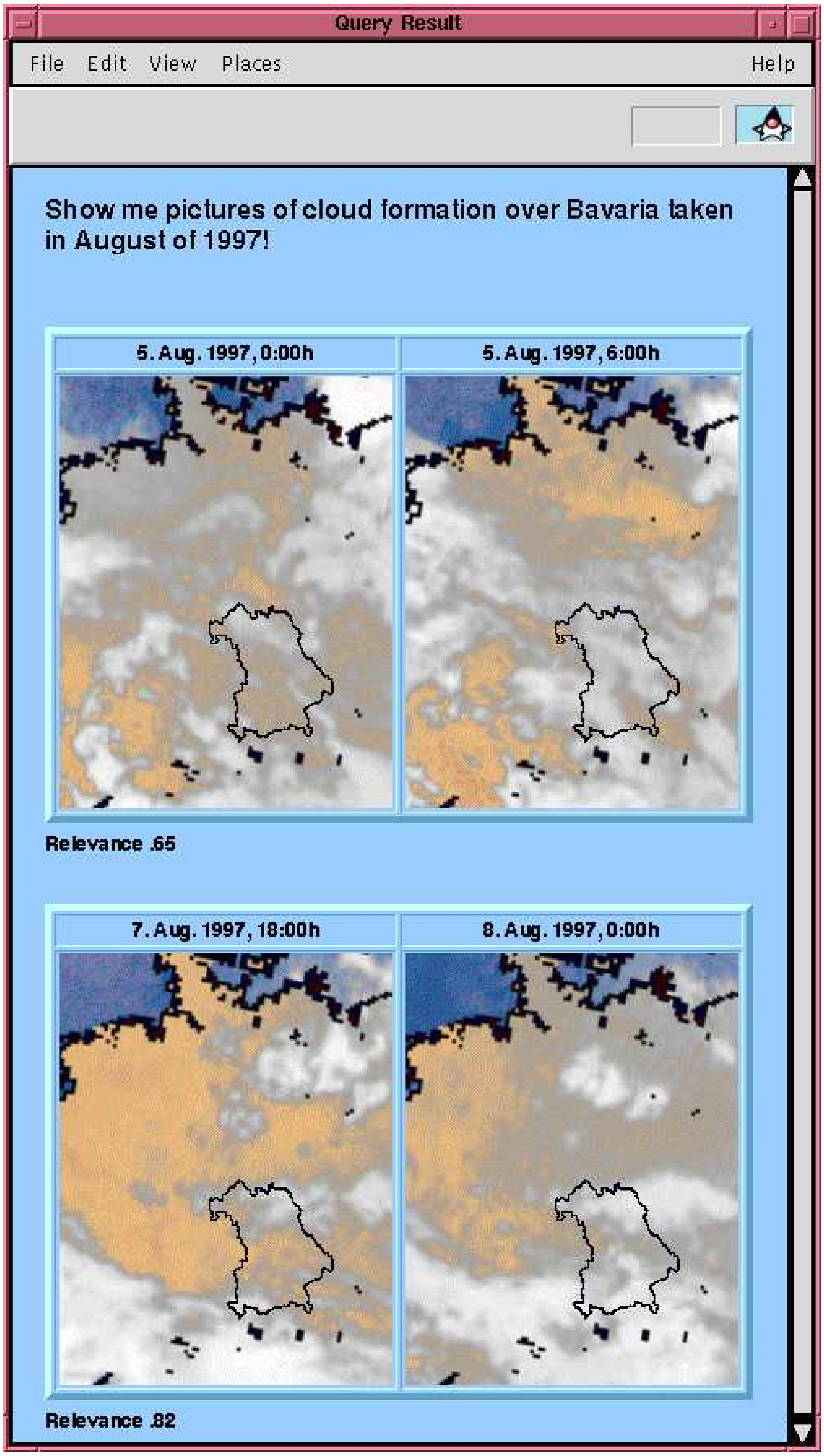

Figure 3: Detected documents (relevant segments)

References

[1] Alois Knoll, Ingo Glöckner, Hermann Helbig, and Sven Hartrumpf. A system for the content-based retrieval of textual and non-textual documents. TR97-03, Technische Fakultät, Universität Bielefeld, November 1997.

[2] J. Biskup, J. Freitag, Y. Karabulut, and B. Sprick. A mediator for multimedia systems. In *Proceedings 3rd International Workshop on Multimedia Information Systems*, Como, Italia, September 1997.

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

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

[5] J. van Benthem. Questions about quantifiers. *Journal of Symbolic Logic*, 49, 1984.

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

[7] L.A. Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Comput. and Math. with Applications*, 9:149–184, 1983.

[8] D. Dubois and H. Prade. *Possibility Theory – An Approach to the Computerized Processing of Uncertainty*. Plenum, New York, 1988.

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

[10] G.J. Klir and T.A. Folger. *Fuzzy Sets, Uncertainty, and Information*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[11] I. Bloch. Information combination operators for data fusion: a comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.

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

[13] Ingo Glöckner. DFS – an axiomatic approach to fuzzy quantification. TR97-06, Technische Fakultät, Universität Bielefeld, November 1997.

[14] Ingo Glöckner. Efficient histogram-based evaluation of fuzzy generalized quantifiers in DFS. Technical report, Technische Fakultät, Universität Bielefeld, 1998. (to appear Jan 1998).

[15] 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.

[16] L.A. Zadeh. Syllogistic reasoning in fuzzy logic and its application to usuality and reasoning with dispositions. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(6):754–763, Nov./Dec. 1985.

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

[18] Dan Ralescu. Cardinality, quantifiers, and the aggregation of fuzzy criteria. *Fuzzy Sets and Systems*, 69:355–365, 1995.

[19] R.R. Yager. Approximate reasoning as a basis for rule-based expert systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 14(4):636–643, Jul./Aug. 1984.

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

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

[22] R.R. Yager. Families of OWA operators. *Fuzzy Sets and Systems*, 59:125–148, 1993.

[23] Fritz Hamm. *Natürlich-sprachliche Quantoren. Modelltheoretische Untersuchungen zu universellen semantischen Beschränkungen (Natural Language Quantifiers. Modell-Theoretical Investigations in Universal Semantical Constraints)*. Number 236 in Linguistische Arbeiten. Niemeyer, Tübingen, 1989.

[24] S. Miyamoto. *Fuzzy Sets in Information Retrieval and Cluster Analysis*. Kluwer Academic Publishers, Dordrecht, 1990.

[25] H.L. Larsen and R.R. Yager. The use of fuzzy relational thesauri for classificatory problem solving in information retrieval and expert systems. *IEEE Transactions on System, Man, and Cybernetics*, 23(1):31–41, Jan./Feb. 1993.

Bisher erschienene Reports an der Technischen Fakultät
Stand: November 1997

96–05 Future University Computing Resources
(Peter Ladkin)

96–06 Lazy Cache Implements Complete Cache
(Peter Ladkin)

96–07 Formal but Lively Buffers in TLA+
(Peter Ladkin)

96–08 The X-31 and A320 Warsaw Crashes: Whodunnit?
(Peter Ladkin)

96–09 Reasons and Causes
(Peter Ladkin)

96–10 Comments on Confusing Conversation at Cali
(Dafydd Gibbon und Peter Ladkin)

96–11 On Needing Models
(Peter Ladkin)

96–12 Formalism Helps in Describing Accidents
(Peter Ladkin)

96–13 Explaining Failure with Tense Logic
(Peter Ladkin)

96–14 Some Dubious Theses in the Tense Logic of Accidents
(Peter Ladkin)

96–15 A Note on a Note on a Lemma of Ladkin
(Peter Ladkin)

96–16 News and Comment on the AeroPeru B757 Accident
(Peter Ladkin)

97–01 Analysing the Cali Accident With a WB-Graph
(Peter Ladkin)

97–02 Divide-and-Conquer Multiple Sequence Alignment
(Jens Stoye)

97–03 A System for the Content-Based Retrieval of Textual and Non-Textual Documents Based on Natural Language Queries
(Alois Knoll, Ingo Glöckner, Hermann Helbig und Sven Hartrumpf)

97-04 Rose: Generating Sequence Families
(Jens Stoye, Dirk Evers und Folker Meyer)

97-05 Fuzzy Quantifiers for Processing Natural Language Queries in Content-Based Multimedia Retrieval Systems
(Ingo Glöckner und Alois Knoll)