Fuzzy modeling in terms of surprise

Arnold Neumaier

Institut für Mathematik, Universität Wien
Strudlhofgasse 4, A-1090 Wien, Austria
email: neum@cma.univie.ac.at
WWW: http://solon.cma.univie.ac.at/~neum/

Abstract. This paper presents a new approach to fuzzy modeling based on the concept of surprise. The new concept is related to the traditional membership function by an antitone transformation. Advantages of the surprise approach include:
1. It has a consistent semantic interpretation.
2. It allows the joint use of quantitative and qualitative knowledge, using simple rules of logic.
3. It is a direct extension of (and allows combination with) the least squares approach to reconciling conflicting approximate numerical data.
4. It is ideally suited to optimization under imprecise or conflicting goals, specified by a combination of soft and hard interval constraints.
5. It gives a straightforward approach to constructing families of functions consistent with fuzzy associative memories as used in fuzzy control, with tuning parameters (reflecting linguistic ambiguity) that can be adapted to available performance data.

1 Introduction

Prologue. Once upon a time, the country of Fuzzyland had a leadership that was famous for its wisdom in governing the welfare of its people. The king spent most of his time traveling through his kingdom to be up to date on the quality of life in his country. Every citizen was entitled and encouraged to present to him and his deputies: complaints about things that didn't go as smoothly as desirable; ideas about improving services, laws, or institutions; observations about faults and failures of public offices; examples of excellence that deserved imitation; and whatever else might benefit the development of the country.

The king listened, judged the quality and value of the cases presented, and then gave orders to his servants, couriers, and ministers. He saw the need for this, the desirability of that, the usefulness of this and the danger of that. And he passed on order by order, as he saw fit. Many different people presented their ideas and concerns, and the king responded with his heart, not caring whether or not his orders were consistent.

Other kingdoms had their rigid laws or self-centered rulers that left many dissatisfied, but Fuzzyland had something special that left the people happier than elsewhere and visitors from everywhere spread its fame. In those times, a king's word was a command that had to be obeyed, and whoever didn't obey risked his life. Plenty were the stories where wise men lost their life or position because they didn't or couldn't satisfy their ruler's many wishes. Not so in Fuzzyland.

> Among the minister's advisors were some mathematicians who worked out a scheme to satisfy their king. The key idea was to translate the fuzzy statements of the king's orders into objective estimates of the surprise he'd feel should he find out how his orders had been followed. Then, equipped with the list of orders of the past, they employed an army of computer slaves to find the least surprising set of actions, and ensured that these actions were carried out properly and efficiently.
>
> Whenever the king was impatient about a request not complied by to the letter, the ministers had data and diagrams to show him that by relaxing just a few of his requirements a little they were able to satisfy many other of his requests to a degree beyond expectation. And the king, whose primary concern was the welfare of the people, was satisfied and blessed his ministers and their wisdom. The country prospered, and the citizens of Fuzzyland blessed their king for his habit to decree the impossible but to be satisfied with the best compromise that could be found.

In this paper we present a new and natural way to formalize qualitative statements about a situation that have an inherent lack of precision. Such statements and whatever is done with them are nowadays usually provided with the label **fuzzy** if it is impractical or impossible to give this lack of precision a quantitative probabilistic interpretation. Two extremes of fuzzy statements are so-called **crisp** statements that do not have any inherent uncertainty, and **interval** statements, where uncertainty is complete except for knowing a range of values. In the latter case, there is usually an associated weaker crisp statement specifying only an inclusion.

In an optimization context, statements expressing constraints on the variables are called **hard constraints** if the statements are crisp, and **soft constraints** if the statements are fuzzy. In particular, **interval constraints** specify that some expressions are limited to intervals, and may be hard constraints if the interval bounds have to be strictly respected, or soft constraints if limited compromises are acceptable. Traditionally, soft interval constraints are modelled by triangular or trapezoidal **fuzzy numbers** [6, 11].

One of the reasons that one wants to be able to make use of qualitative statements is that they are quite abundant in many contexts, and convey useful information that may be decisive for getting an adequate solution of a practical problem.

Another reason is that in many situations we only have this kind of vague and sometimes even conflicting knowledge, and still want to be able to draw useful consequences. This is especially relevant for problems in artificial intelligence, where it is important to be able to reason under uncertainty.

Finally, there are many problems whose solution only need qualitative information that is easy to pick up, while obtaining precise information to set up an accurate mathematical model may be costly or even impossible. This is frequently the case for problems of control design in an uncertain environment [4, 12, 17, 22], and has lead to many innovative engineering applications, such as in traffic control [10], or the fuzzy control units in cameras and vacuum cleaners.

In the following, we discuss fuzzy modeling based on the concept of (expected) **surprise**, see Section 3. This deviates from tradition [7], where the concept of **possibil-**

ity is the basic notion. The two concepts are related by an antitone transformation that appears in traditional treatments in the context of Archimedian t-norms; see Section 5.

The surprise approach has several advantages. It has a well-defined semantic interpretation with less inherent ambiguity than possibility (cf. Remark 5.1), that is especially adapted to the precise formulation of imprecise goals. (It is less suitable for modeling uncertainty that reflects various degrees of optimism, and for reasoning with uncertain facts.) It allows the combination of quantitative and qualitative knowledge in a natural way. Indeed, see (33), it is a direct extension of the time-honored least squares approach to reconciling conflicting approximate numerical data; see Example 2.1.

In particular, it gives a sound methodology for interpreting **optimization problems** involving imprecise goals containing hard and soft interval constraints (Section 7), and for specifying smooth functions in terms of a **fuzzy associate memory** (Section 8). An interesting application to radiation therapy planning will be given elsewhere (LODWICK et al. [14]).

The paper is completely self-contained; we therefore give references only in a few places where direct contact is made with specific results of others. However, the material here owes its existence to a large body of previous work on fuzzy modeling that inspired me to develop the present approach. See, e.g., [6, 11] for more on fuzzy modeling, [16] for more about interval analysis, and [2, 13, 18, 27] for more on fuzzy optimization.

A referee pointed out that a theory of possibility in terms of surprise was already considered by SHACKLE [20]. While the intuition presented there is very close to that of the present paper (and I recommend reading his Chapter IX), the resulting formalization is completely different. In particular, his surprise is $[0,1]$-valued (while ours is $[0,\infty]$-valued) with $s \vee \neg s = 0$ universally (while ours satisfies this only for crisp s). This points to a difficulty with Shackle's surprise in an example discussed by MCCAIN [15]. Moreover, Shackle's surprise has the nondifferentiable rule $s_1 \vee s_2 = \min(s_1, s_2)$, while we have the differentiable rule (13) below.

SPOHN [23, 24] and SHENOY [21] also use surprise-like functions, valued on the integers, though. Their emphasis is on reasoning with (or change of) beliefs as uncertain facts, while we are interested primarily in formalizing uncertain goals. Another fuzzy theory based (as the present surprise approach) on the scale $[0,\infty]$ but differing in the axioms was proposed by DUBOIS & PRADE [8] for so-called toll sets.

None of these earlier approaches satisfies (as the present approach) $\neg\neg s = s$, and there are no connections with least squares, which is the basic motivation for the present approach. The most conspicuous formal difference seems to be that the older approaches all have a set-function based setting, while (see Section 6) such a setting is unnatural in the present approach.

2 Fuzzy statements

Our first task is to explain what is meant by a fuzzy statement and how to relate with it a surprise value.

As a motivating example, suppose that someone says,

$$\text{“I'll come tomorrow afternoon around 4 o'clock.”} \tag{1}$$

That this is a fuzzy statement is signified by the modifier **around** that makes the following time specification a little vague. The corresponding non-fuzzy (or crisp) statement

$$\text{“I'll come tomorrow afternoon at 4 o'clock”} \tag{2}$$

specifies a precise arrival time 4:00pm. But from our knowledge of life we know that even the statement (2) should get a fuzzy interpretation since it probably implies an arrival time in a small interval around 4:00pm, though closer to the target time than implied in (1). On the other hand, the statement

$$\text{“I'll come tomorrow afternoon”} \tag{3}$$

is much more uncertain, allowing the arrival time to vary in a wide interval. Uncertainty of a different kind is introduced by the statement

$$\text{“I'll probably come tomorrow afternoon around 4 o'clock.”} \tag{4}$$

The modifier *probably* makes the arrival time to some degree unpredictable, though the expectation to arrive would be large near 4:00pm.

How do we assess the meaning of (1) and (2)? If we know the person making the statement well enough, we know from past observation how closely his language corresponds to reality and will infer from this the amount of uncertainty in the statement. If we don't know the person, we replace our trust in past observations by trust in established traditions or in our general experience of the reliability of language, and infer from it the contents of the statement.

Inversely related to the amount of uncertainty is the surprise we feel when the person arrives at 3:25pm, say. We would find it perfectly natural in case (3), be probably a little surprised about the early arrival given (1), or (4), and (assuming the person is usually quite on time) would be really surprised given (2). On the other hand, if the person arrives at 8:00pm we would be very surprised in cases (1)–(3), but less so in case (4).

From this example we may deduce some general features of fuzzy statements.

(i) Each fuzzy statement involves one or several variables to which the statement applies, in our case the *arrival time*. The variable(s) may appear in the statement explicitly, such as the *weight* in

$$\text{“the weight is very high”} \tag{5}$$

or less explicit as in (1)–(4), or completely implicit, as in

$$\text{“the weather is fine”} \tag{6}$$

which may involve information about temperature, humidity, cloudiness, etc.

(ii) The interpretation of the statement (if not already the variables) is context dependent, here dependent on the assumed knowledge about the person. Assessment of the information contents may be based either on previous data about analogous situations, or on default model assumptions.

(iii) One or several **modifiers**, such as *very, probably, perhaps, hardly, extremely*, may change the implied uncertainty of the information provided (cf. SCHMUCKER [19]).

(iv) Different statements about the same situation may be discriminated on the basis of the amount of surprise caused by the disclosure of the truth.

(v) While one can clearly distinguish between no surprise, little surprise, and high surprise, it is frequently difficult or a matter of subjective judgment to specify the amount of surprise quantitatively. (As the language shows, the statements about surprise look themselves fuzzy.)

Fuzzy modeling, or more precisely surprise modeling, consists in fixing quantitatively the amount $s(x|E)$ of surprise in disclosing the value of a variable (or vector) x given a statement E. In short, $s(x|E)$ is the **surprise** of x given E. The analysis of a fuzzy model in which surprise functions of a number of statements are given then consists in drawing useful consequences for the application in question. As remark (v) above indicates, the modeling stage is to some extent subjective, and this is reflected in the fact that the surprise functions produced contain a number of adjustable tuning parameters (see Section 3). These must be tuned with suitable data to provide optimal performance in the applications.

The **most plausible** values are those values of x for which $s(x|E)$ is minimal, i.e., which are least surprising given E. Note that we shall allow the minimal value to be positive, as an indication that the information given in E is to some extent inconsistent. An example of this is a composite statement that simultaneously claims “x is approximately 3.5 and x is approximately 3.6”. Highly contradictory statements such as “x is probably small and x is probably large” will be considered as nearly false and hence get a high surprise value $s(x|E)$ for all x.

The fact that such seemingly inconsistent statements can be harmonized is one of the strengths of fuzzy modeling when applied to uncertain reasoning in artificial intelligence, where often different experts provide different, and partly inconsistent, inputs to the information data base.

2.1 Example. The oldest and time-approved method for harmonizing uncertain data is the *least squares method*. In the simplest case, we have m independent measurements of a quantity x with equal accuracy σ, giving rise to fuzzy statements $x \approx x_i \pm \sigma$ $(i = 1 : m)$. If we use the surprise function $s(x|x \approx x_i \pm \sigma) = (\frac{x-x_i}{\sigma})^2$ and assume that the surprise of independent measurements is additive, the total surprise given all m statements about x is $s(x) = \sum(\frac{x-x_i}{\sigma})^2$, and the most plausible value turns out to be the mean $x = \frac{1}{m}\sum x_i$.

This example can be generalized to the multivariate case, by defining the surprise of the statement x =rand(μ, C), that x is an uncertain vector with mean μ and (symmetric, positive definite) covariance matrix C, to be

$$s(x|x = \mathrm{rand}(\mu, C)) = (x - \mu)^T C^{-1}(x - \mu).$$

Note that a reference to a specific probability distribution is neither made nor necessary; in particular, this is an adequate description of many situations where specifying a Gaussian density would (with high confidence) contradict the data available.

3 Fuzzy semantics

The applicability of fuzzy modeling to complex situations is based on the fact that often complex situations can be analyzed in terms of a lot of simple statements involving only a few variables. We agree on the axiom that for two vectors x, x' of variables,

$$s(x|E) = s(x'|E)$$

whenever x and x' contain the same information about the terms referred to in E. This is indeed natural since the surprise should not depend on additional information supplied in x or x' that is not referred to in E. Thus if E is the statement "x_i is tiny", the surprise function, though formally dependent on a vector x, will be a function of x_i alone, i.e., we have

$$s(x|E) = s(x_i|E).$$

Together with logical rules for combining statements by "and", "or" and "not" (see Section 4), and modification rules for the interpretation of modifier such as "probably", "perhaps", etc., we can evaluate the surprise values of complex scenarios and find the most plausible scenario by minimizing the surprise value with respect to the scenario variables.

If a statement E is true for a particular value of x then there is no surprise,

$$s(x|E) = 0 \quad \text{if } E(x) \text{ is true.} \tag{7}$$

And if E is false then the occurrence of x, being impossible, would be infinitely surprising,

$$s(x|E) = \infty \quad \text{if } E(x) \text{ is false.} \tag{8}$$

These are the only possibilities for **crisp** statements, i.e., unambiguous logical propositions. For general fuzzy statements, of course, any value in between may be possible when $E(x)$ is vague, i.e., neither true nor false. Thus we take the set $[0,\infty] = \mathbb{R}_+ \cup \{\infty\}$ as the set of possible surprise values, with 0 representing **true** and ∞ representing **false**. Thus surprise can also be interpreted as an "amount of falseness" (or lack of appropriateness) of statement E after knowing x.

To give the surprise values a definite scale we require that the condition

$$s(x_a|E) = 1 \tag{9}$$

signifies threshold values where the appropriateness of E or its negation $\neg E$ given $x = x_a$ is judged to be completely ambivalent, without preference for either. We may call the solutions x_a of (9) the **ambivalence points** of E. The set of points with $s(x|E) \leq 1$ will be called the **preference region** of E. Points $\hat{x}$ minimizing $s(x|E)$ also have special significance as the least surprising and hence **most plausible points** for a statement E. (Compare this with KOSKO's [12] notion of points of maximum entropy.)

While fuzzy logic (discussed in the next section) can be used to assign surprise values to composite statements, the surprise values of the constituents must be determined by some other means. Before discussing linguistic statements we shall look at a variety of **objective fuzzy statements** in a single variable x with precisely defined surprise values.

case	statement "x is E"	surprise $s(x\|E)$	restrictions
1	$x \approx \alpha \pm \sigma$	$(\frac{x-\alpha}{\sigma})^2$	$\sigma > 0$
2	$x \approx< \alpha \pm \sigma$	$(\frac{\sigma-\alpha+x}{2\sigma})_+^2$	$\sigma > 0$
3	$x \approx> \alpha \pm \sigma$	$(\frac{\sigma+\alpha-x}{2\sigma})_+^2$	$\sigma > 0$
4	$x \approx\in [\alpha,\beta] \pm \sigma$	$\frac{(\alpha-x)_+^2+(x-\beta)_+^2}{\sigma^2}$	$\alpha \leq \beta,\ \sigma > 0$
5	$x \approx> \alpha,\ x > \gamma$	$\frac{\alpha-\gamma}{(x-\gamma)_+}$	$\gamma < \alpha$
6	$x \approx< \beta,\ x < \delta$	$\frac{\delta-\beta}{(\delta-x)_+}$	$\beta < \delta$
7	$x \approx> [\gamma,\alpha]$	$(\frac{(\alpha-x)_+}{(x-\gamma)_+})^2$	$\gamma < \alpha$
8	$x \approx< [\beta,\delta]$	$(\frac{(x-\beta)_+}{(\delta-x)_+})^2$	$\beta < \delta$
9	$x \approx \alpha,\ \gamma < x < \delta$	$(\frac{(\alpha-x)_+}{(x-\gamma)_+})^2 + (\frac{(x-\alpha)_+}{(\delta-x)_+})^2$	$\gamma < \alpha < \delta$
10	$x \approx\in [\alpha,\beta],\ \gamma < x < \delta$	$(\frac{(\alpha-x)_+}{(x-\gamma)_+})^2 + (\frac{(x-\beta)_+}{(\delta-x)_+})^2$	$\gamma < \alpha \leq \beta < \delta$

Table 1: Some fuzzy statements and their surprise values

In the statements defined in Table 1, the surprise function is taken as a simple differentiable truncated rational consistent with the intuitive interpretation, and with surprise values of zero, one and ∞ at the natural places; we write $x_+ = \max(x, 0)$. Note that Case 1 is a special case of Case 4, and Case 9 of case 10.

We now discuss the use of objective fuzzy statements to model linguistic statements. Since there is some ambiguity and subjectivity (or at least considerable dependence on the context) in linguistic statements like "x is large", the latter must be replaced by objective fuzzy statements. It is usually easy to select for a given primitive linguistic statement an objective fuzzy statement from Table 1 matching its intuitive contents.

Note that the statements defined by Table 1 all involve some **tuning parameters** α, β, etc.. Typically, the context of a linguistic description only gives a crude guide for their values. However, they can be tuned to the particular context in which the ambiguity is resolved. To decide upon the right tuning parameters is often a matter of judgment, and part of the modeling process. In many cases, the tuning parameters can be determined by a training procedure on performance data; cf. Section 8. In other cases, one may use the relation to fuzzy numbers discussed in Section 5 to specify an objective fuzzy statement.

Of course, we want to have identical linguistic statements occurring in equivalent contexts modeled by objective fuzzy statements with identical tuning parameters. This generally limits the total number of tuning parameters to be specified or estimated to a fairly small number.

4 Fuzzy logic

Our next task is to give the concept of surprise some depth by relating it to logic. A few simple requirements *uniquely* determine the logical operations for surprise values.

Clearly, we need to define the logical operations on $[0, \infty]$ in such a way that the restriction to the set $\{0, \infty\} = \{$**true, false**$\}$ recovers standard Boolean algebra. For the **and**-operation $\wedge$, consistency with the least squares method, necessary to combine qualitative and numerical uncertain knowledge, requires (cf. Example 2.1) that we define

$$s_1 \wedge s_2 = s_1 + s_2. \tag{10}$$

Table 1 is consistent with this definition since

$$(x \approx> [\gamma, \alpha]) \wedge (x \approx< [\beta, \delta]) = \begin{cases} (x \approx\in [\alpha, \beta], \gamma < x < \delta) & \text{if } \alpha \leq \beta, \\ (x \approx \alpha, \gamma < x < \delta) & \text{if } \alpha = \beta. \end{cases}$$

To find the appropriate definition for the **not**-operation $\neg$ we note that the semantics requires that $\neg s$ is an antitone function of s satisfying $\neg\neg s = s$. Moreover, the ambivalent surprise value $s = 1$ must be the only value with $\neg s = s$. Later applications also require smoothness. This leaves as only natural choice the definition

$$\neg s = s^{-1}. \tag{11}$$

Table 1 is also consistent with this definition since

$$\neg(x \approx> [\gamma, \alpha]) = (x \approx< [\gamma, \alpha]).$$

The formula for the **or**-operation $\vee$ is determined by the axiom

$$\neg(s_1 \vee s_2) = (\neg s_1) \wedge (\neg s_2). \tag{12}$$

Indeed, this implies

$$s_1 \vee s_2 = \neg((\neg s_1) \wedge (\neg s_2)) = (s_1^{-1} + s_2^{-1})^{-1} = s_1 s_2/(s_1 + s_2). \tag{13}$$

Similarly, the value of the implication $\Rightarrow$ is determined by the formula

$$(s_1 \Rightarrow s_2) := \neg(s_1 \wedge \neg s_2) = \frac{1}{s_1 + s_2^{-1}} = s_2/(1 + s_1 s_2). \tag{14}$$

(In these formulas, ∞ must be treated as a limit of huge numbers.) It is easy to see that $\wedge, \vee$ are commutative and associative, with

$$s_1 \wedge \ldots \wedge s_n = \sum_{i=1}^{n} s_i, \tag{15}$$

$$s_1 \vee \ldots \vee s_n = 1 \Big/ \sum_{i=1}^{n} s_i^{-1}. \tag{16}$$

Moreover,

$$s_2 = \neg s_1 \iff s_1 = \neg s_2,$$

$$(\neg s_1 \Rightarrow \neg s_2) = (s_2 \Rightarrow s_1).$$

Equipped with the logical operations, we require the surprise function to satisfy the following three axioms:

$$s(x|E) \in [0, \infty], \tag{17}$$

$$s(x|E_1 \wedge E_2) = s(x|E_1) \wedge s(x|E_2), \tag{18}$$

$$s(x|\neg E) = \neg s(x|E). \tag{19}$$

It is straightforward to deduce as consequences of the axioms (17)–(19) the formulas

$$s(x|E_1 \wedge \ldots \wedge E_n) = s(x|E_1) \wedge \ldots \wedge s(x|E_n) = \sum_{i=1}^{n} s(x|E_i), \tag{20}$$

$$s(x|E_1 \vee \ldots \vee E_n) = s(x|E_1) \vee \ldots \vee s(x|E_n) = 1 \Big/ \sum_{i=1}^{n} s(x|E_i)^{-1}, \tag{21}$$

$$s(x|E_1 \Rightarrow E_2) = (s(x|E_1) \Rightarrow s(x|E_2)) = \frac{s(x|E_2)}{1 + s(x|E_1)s(x|E_2)}. \tag{22}$$

5 Fuzzy numbers and interval constraints

Using $0 =$ **true** and $\infty =$ **false** as truth values is nonstandard, and one can also develop fuzzy logic based on the standard truth values $0 =$ **false** and $1 =$ **true**. This is the traditional way taken in fuzzy set theory, where a **fuzzy set** is defined by a membership function with values $d(x) \in [0, 1]$. $d(x)$ is called the **possibility** of x; if it is certain that x belongs (resp. does not belong) to the set then $d(x) = 1$ (resp. $d(x) = 0$).

To connect the fuzzy set approach with surprise, we quantify the evidence for a value x given E by a **degree of consistency** $d(x|E) \in [0, 1]$, with $d(x|E) = 0$ if x is impossible given E, and $d(x|E) = 1$ if x is fully consistent with E. We may translate any surprise value s into a degree of consistency d by means of any strictly monotone decreasing transformation from $[0, \infty]$ onto $[0, 1]$. The most natural way to do so is the transformation

$$d = \frac{1}{s^e + 1} \tag{23}$$

for some exponent $e > 0$. It maps 0 to 1, ∞ to 0 and yields possibilities such that

$$d(x|\neg E) = 1 - d(x|E).$$

In particular, ambivalence points with surprise value $s = 1$ correspond to points with degree of consistency $d = \frac{1}{2}$. Conversely, one may transform given degrees of consistency to surprise values by the inverse of (23),

$$s = (d^{-1} - 1)^{1/e}. \tag{24}$$

Thus every fuzzy set can be given a surprise interpretation, and indeed infinitely many. (This makes the fuzzy set approach more ambiguous than the surprise approach, and may be a reason for the continuing interpretation debate [5]). Although $e = 1$ seems to be simpler (and has been used in a related context by ZIMMERMANN [25, p.204]), we shall see in a moment that the most useful and hence recommended choice of the exponent is $e = \frac{1}{2}$.

Associated with each statement E is a fuzzy set whose membership function is given by $d(x) = d(x|E)$. Thus possibilities d of a fuzzy set are related to surprise values by (23)–(24), too. In applications, the most prominent fuzzy sets encountered are triangular and trapezoidal fuzzy numbers.

A **trapezoidal fuzzy number** is a fuzzy set such that, for suitable intervals $[\alpha, \beta]$ (the **core**) and $[\gamma, \delta]$ (the **support**, containing the core in its interior), the possibility of a number x is defined by

$$d(x) = \begin{cases} (x-\gamma)/(\alpha-\gamma) & \text{if } x \in [\gamma, \alpha], \\ 1 & \text{if } x \in [\alpha, \beta], \\ (\delta - x)/(\delta - \beta) & \text{if } x \in [\beta, \delta], \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

A **triangular fuzzy number** is a trapezoidal fuzzy number with $\alpha = \beta$, so that the core contains a single number only. Triangular and trapezoidal numbers arise naturally in the modeling of soft interval constraints. Indeed, considered as a fuzzy number, an interval is the limit $[\gamma, \delta] = [\alpha, \beta]$ of a trapezoidal number, in which support and core agree. Therefore a hard interval constraint $f(x) \in [\alpha, \beta]$ may be relaxed to a soft interval constraint by allowing that the bounds need to be satisfied only to some degree, specified by the possibility $d(x)$ of a trapezoidal number with core $[\alpha, \beta]$. Alternatively, a hard open interval constraint $f(x) \in\ ]\gamma, \delta[$ may be strengthened by discouraging values of $f(x)$ that are too close to the bounds, by requiring that $f(x)$ is close to some target value α, to some degree specified by the possibility $d(x)$ of a triangular number with support $[\gamma, \delta]$.

In both cases, the possibility has a natural interpretation as the degree of consistency with a fuzzy statement, namely $f(x) \approx\in [\alpha, \beta]$, $\gamma < x < \delta$ in the first case, and $\gamma < x < \delta$, $x \approx \alpha$ in the second case. And indeed, the corresponding surprise functions from Table 1 are obtained from the degree of consistency (25) for trapezoidal (or triangular) numbers using the correspondence rules (23)–(24) with $e = \frac{1}{2}$. (Note that, for $e = 1$, nondifferentiable functions would arise, a significant disadvantage in the applications to fuzzy optimization and fuzzy control.)

5.1 Remark. We may use (23) and (24) to translate the logical operations on surprise values to logical operations involving possibilities. The resulting formulas correspond precisely to the operations defined by DOMBI's [3] one-parameter family of t-norms and co-t-norms. Conversely, if we have a fuzzy logic over $[0, 1]$ in which $\wedge$ is given by an Archimedean t-norm, so that

$$d(x|E \wedge E') = f^{(-1)}(f(d(x|E)) + f(d(x|E')))$$

for some strictly decreasing, surjective function $f : [0, 1] \to [0, \infty]$ with inverse $f^{(-1)}$ (see [11, Theorem 3.11]), then $s = f(d)/f(\frac{1}{2})$ is the natural (additive and correctly normalized) surprise value associated with the possibility d. Thus a single surprise calculus is equivalent with an arbitrary Archimedean fuzzy logic calculus. In particular, the ambiguity in the rules for traditional fuzzy logic, which has lead to continuing controversial views regarding a consistent semantical interpretation (see, e.g., [5, Section 1.6]) is eliminated in a natural way with the present surprise interpretation.

6 Dependence and modifiers

Unfortunately, the rules (10)–(14) do not satisfy all the properties used to from classical logic; most of the inference rules only have very weak substitutes. For example, we have in place of the law of contradiction and the law of excluded middle only

$$s \wedge (\neg s) \geq 2 \quad s \vee (\neg s) \leq \tfrac{1}{2},$$

since $s \wedge (\neg s) = s + s^{-1} = 2 + (s-1)^2/s \geq 2$ and $s \vee (\neg s) = ss^{-1}/(s + s^{-1}) = s/(2s + (s-1)^2) \leq \frac{1}{2}$. This still looks reasonable for fuzzy statements. At first sight less intuitively, we also have

$$s \wedge s = 2s \neq s.$$

However, this has the natural interpretation that repetition of a statement emphasizes it and hence makes exceptions more surprising. It is indeed more surprising that two independent surprising events of the same kind happen than if only one happens. But if the two events are identical, the formula overestimates the surprise.

Thus, in analogy to *interval arithmetic* [16], which is another way to represent incomplete knowledge, fuzzy logic reflects an inherent independence assumption in different occurrences of the same variable. The surprise calculus cannot differentiate between apparent 'confirmation' of a statement by independent means and phony repetition of an opinion to make it sound stronger. This susceptibility is a feature due to incomplete information that automatic reasoning devices will have to share with human reasoning, unless the dependence is recognized and explicitly modeled with appropriate multivariate surprise functions. (However, a discussion of this is beyond the scope of the present paper.)

The dependence problem also implies that it is impossible to give a consistent set function version of the surprise calculus. Indeed, in analogy to measure theory, one would like to introduce (for a variable or vector x with fixed but unknown – and perhaps context dependent – value in some domain Ω, and arbitrary subsets A of Ω) the set function $s(A) := s(x|x \in A)$. Then $s(A \cap B) = s(x|x \in A \wedge x \in B) = s(x|x \in A) + s(x|x \in B) = s(A) + s(B)$, but this would give for $A = B$ the conclusion $s(A) = 2s(A)$, which forces $s(A) \in \{0, \infty\}$. In particular, this calculation teaches us that asserting "$x \in A$" consistent with the rules of set theory necessarily implies that this is a crisp statement; a corresponding fuzzy statement "$x \approx\in A$" with finite, nonzero surprise values cannot satisfy the rule $x \in A \cap B \Leftrightarrow x \in A \wedge x \in B$ implicitly used in the above argument.

(The missing set function version also distinguishes the present approach from the surprise-like approaches mentioned in the introduction. Of course, using the results of the previous section, one may reinterpret the standard fuzzy set formulation in surprise terms, and then invert the interpretation. This amounts to defining, for a fixed statement E, the **possibility** $P(A) := \sup\{d(x|E) \mid x \in A\}$ of a set A, where $d(x|E)$ is the degree of consistency (23). Then $P(A)$ satisfies the traditional possibility axioms [7]. However, such a statement dependent definition looks rather unnatural.)

The lack of strong valid implications makes arguing with fuzzy logic weak since the (worst case, possibilistic) logic has to account for that part of the logical chain that happens to fail just for the case in question. Therefore, *reasoning with uncertain facts* is done more consistently using causal (Bayesian belief) networks [1].

However, the surprise approach is ideally suited to expressing the logical relations in

modifier	β	$\pm$
extremely	$\gg 1$	$+$
very	> 1	$+$
probably	< 1	$+$
perhaps	$\ll 1$	$+$
hardly	< 1	$-$
not	$= 1$	$-$

Table 2: Semantic interpretation of some modifiers

uncertain goals, where probability assignments are often meaningless. We show this below by looking at fuzzy optimization and fuzzy function specification.

On the other hand, this repetition phenomenon gives us a clue how to handle modifiers. Indeed, repetition of a statement (resulting in multiplication of the surprise function by 2) can be regarded as an emphasis of the same sort as adding a modifier like *very*, which therefore should be interpreted as multiplication by some (context dependent) number > 1. Similarly, deemphasizing modifiers like *probably* can be interpreted as multiplication by some (context dependent) number < 1. Since there are also modifiers such as *hardly* which are softened versions of negation, we take the general modifier to be defined by the semantics

$$s(x|\textit{modifier } E) = \beta s(x|E)^{\pm 1}. \tag{26}$$

Table 2 gives the magnitude of β and the sign of the exponent suitable for some modifiers.

In the present rules for modifiers, β is a tuning parameter whose precise value must be found out from the context, as for the tuning parameters in the primitives discussed in Section 3. This contrasts with the rigid definition of modifiers in ZIMMERMANN [25, pp.238–241]. Note that translation of his modifier rules into surprise values does *not* give rules of the form (26).

7 Fuzzy optimization

One of the important areas where fuzzy techniques may be very useful is optimal planning in the presence of imprecise goals and constraints. For example, there has been quite a lot of work on fuzzy scheduling and on fuzzy optimization [2, 13, 18, 27].

As suggested by ZIMMERMANN [27], we assume that if there is an objective function, it is rewritten as a fuzzy constraint that expresses the degree to which the fuzzy objective is satisfied. In particular, this allows the formulation of several, possibly conflicting and possibly imprecise goals; each goal becomes a fuzzy constraint.

As the Prologue indicated, a few, but not too many, items of low consistency with the goals are acceptable if they make most other items consistent with the goals. Wanted is the most acceptable compromise. In terms of surprise, we want to find the most plausible combination of parameters consistent with the fuzzy goals and satisfying the hard constraints. In accordance with our development so far, the most plausible combination of parameters is the one that minimizes the total surprise.

In order to have a clear linguistic interface for the fuzzy part of an optimization problem, we describe the latter in terms of so-called fuzzy scales. On the linguistic level, a **fuzzy scale** is a finite set U of **descriptors** that specify fuzzy properties of a single real variable, roughly dividing the real line, or a subinterval of interest, into overlapping regions corresponding to each descriptor. Frequently, but not always, the scale has an intrinsic ordering. Typical examples of fuzzy scales are

- *tiny, very small, small, medium size, large, very large, huge,*
- *large negative, small negative, roughly zero, small positive, large positive,*
- *never, almost never, rarely, seldom, occasionally, sometimes, frequently, often, very often, almost always, always,*
- *none, hardly any, few, several, some, a good fraction of, many, most, almost all, all,*

and their subsets. Associated with each of the fuzzy descriptors u in a fuzzy scale U are fuzzy statements of the form $E =$ "x is u", and on the semantic level, the fuzzy scale gets an objective interpretation by defining the surprise values

$$s(x|x \text{ is } u) := s_u(x) \quad \text{for } u \in U,$$

in a way consistent with the intuitive meaning of the descriptors u.

Now we are ready to formulate in precise mathematical terms a large and useful class of fuzzy optimization problems. We consider a **fuzzy optimization problem** of the form

$$\begin{array}{ll} \min & \text{total surprise} \\ \text{s.t.} & x \in \mathbf{x}, \quad F(x,y) \in \mathbf{F}, \\ & y_i \text{ is } u_i \quad (i = 1:d). \end{array} \tag{27}$$

Here s.t. is short for "subject to" (the list of constraints following),

$$\mathbf{x} = [\underline{x}, \bar{x}] = \{x \in \mathbb{R}^n \mid \underline{x} \le x \le \bar{x}\}$$

(where $\underline{x} \in (\mathbb{R} \cup \{-\infty\})^n$, $\bar{x} \in (\mathbb{R} \cup \{\infty\})^n$, $\underline{x} \le \bar{x}$, and inequalities between vectors are interpreted componentwise) is a bounded or unbounded **box** in $\mathbb{R}^n$ bounding the crisp variables x_k $(k = 1:n)$; $F(x,y)$ is a vector of m continuous constraint functions $F_1(x,y), \dots, F_m(x,y)$ with values contained in the box $\mathbf{F}$; and u_i are descriptors from fuzzy scales that describe the properties of the fuzzy quantities y_i. Usually, the fuzzy quantities occur in groups having a similar semantics and hence using descriptors from the same fuzzy scale; thus only one or a few such scales need to be defined. To translate a fuzzy optimization problem into a precise mathematical statement we need to give the fuzzy scales objective meanings by defining for each descriptor $u \in \{u_1, \ldots, u_d\}$ surprise values

$$s(y \text{ is } u) := s(y|u), \tag{28}$$

where for each u, $s(y|u)$ is a specific function derived from the surprise calculus already discussed. The total surprise is then given by

$$s_{\text{total}} = \sum_{i=1}^{d} s(y_i|u_i),$$

and the fuzzy optimization problem (27) translates into a standard nonlinear program of the form

$$\begin{array}{ll} \min & \sum_{i=1}^{d} s(y_i|u_i) \\ \text{s.t.} & x \in \mathbf{x}, \quad F(x,y) \in \mathbf{F}. \end{array} \tag{29}$$

7.1 Remarks. (i) Note that the semantics of the fuzzy optimization problem (27) is completely specified by the objective fuzzy statements (28) defining the fuzzy scales used. Thus only a single interpretation stage is needed to translate a problem given in terms of linguistic information for the fuzzy part into a precise and well-defined mathematical task consistent with the linguistic usage. This is in contrast to traditional fuzzy optimization where, in addition to the stage giving a precise definition of the fuzzy statements, a second interpretation stage is needed that defines what should be meant by a solution of the associated fuzzy optimization problem. And indeed, different proposals lead to different answers. This shows that – unlike stochastic modeling and the present surprise modeling – traditional fuzzy modeling of an optimization problem does not define a well-posed mathematical problem.

(ii) Except in objective fuzzy statements such as "$y \approx \alpha \pm \beta$", where the semantics of each fuzzy quantity is defined, fuzzy quantities behave exactly like ordinary numbers. In particular, operations between fuzzy quantities are *ordinary* arithmetical operations. This allows very free modeling and avoids dependence problems to a large extent.

(iii) In principle, the tuning parameters in an objective fuzzy statement may themselves be constrained, or be subject to fuzzy statements, giving a 'Bayesian' view for modeling fuzzy optimization problems.

We now consider the special case where the fuzzy variables only occur as right hand side of the constraints. We take the fuzzy optimization problem to have the special form

$$\begin{array}{ll} \min & \text{total surprise} \\ \text{s.t.} & x \in \mathbf{x}, \quad F(x) = y, \\ & y_i \text{ is } u_i \quad (i = 1:d). \end{array} \tag{30}$$

(Inequality constraints with fuzzy right-hand sides can be transformed to equality constraints by modifying the interpretation of the right hand side. Indeed, an inequality constraint $F_i(x) \le y_i'$ (y_i' is u_i') can be rewritten in the equivalent form $F(x) = y$ (y_i is u_i), by introducing the new fuzzy number y_i with surprise $s(y_i|u_i) = \min_{y_i' \ge y_i} s(y_i'|u_i')$; two-sided constraints can be rewritten in a similar way.)

The corresponding nonlinear program

$$\begin{array}{ll} \min & \sum_{i=1}^{d} s(y_i|u_i) \\ \text{s.t.} & x \in \mathbf{x}, \quad F(x) = y \end{array} \tag{31}$$

is clearly equivalent with the bound constrained problem

$$\begin{array}{ll} \min & \sum_{i=1}^{d} s(F_i(x)|u_i) \\ \text{s.t.} & x \in \mathbf{x}. \end{array} \tag{32}$$

This formulation applies to many fuzzy scheduling problems; the difficulty of scheduling problems [17, 22] is reflected in the fact that, in these cases, (32) is a global optimization problem with typically many local minimizers.

On the other hand, for **fuzzy linear programs**, where (30) holds with $F(x) = Ax$, and the surprise functions from Table 1, the objective function is smooth and convex, hence (32) is a very well-behaved bound constrained optimization problem.

Moreover, for fuzzy statements of the form $y_i \approx F_i(x) \pm \sigma_i$, (32) reduces to a weighted least squares problem

$$\begin{array}{ll} \min & \sum_{i=1}^{d} \Big(\frac{F_i(x) - y_i}{\sigma_i}\Big)^2, \\ \text{s.t.} & x \in \mathbf{x}. \end{array} \tag{33}$$

in accordance with how tradition treats such statements for a long time.

If hard constraints are present, some of the right hand sides y_i are in fact crisp numbers or intervals, and the corresponding surprise functions take only the values 0 or ∞. Instead of taking the corresponding terms into the objective function, one rather keeps them as hard constraints.

8 Fuzzy function specification

As a second application, we discuss a flexible way of giving a qualitative specification of functions. Input and output variables are again described in terms of fuzzy scales. Suppose that we have $n + 1$ real variables $x_1, \ldots, x_n, y$, and associated fuzzy scales $U_1, \ldots, U_n, U$. These scales may be identical, but they need not be. A **fuzzy associative memory (FAM)** [12] is a list of rules of the form:

$$E(u_1, \ldots, u_n, u): \quad x_1 \,\mathrm{is}\, u_1 \ \wedge \ldots \wedge \ x_n \,\mathrm{is}\, u_n \ \Rightarrow \ y \,\mathrm{is}\, u. \tag{34}$$

We permit multiple rules with the same predicates but different conclusions, to allow the modeling of properties of a function by different, possibly disagreeing experts. The possible values of u_i (and u) are the descriptors from U_i (and U), or the descriptor `any`. The descriptor `any` is used to provide default rules for the case where no information is available about one or several input arguments; the complete lack of information is interpreted as zero surprise, no matter what the actual value of the variable.

Most typically, FAMs are used with $n = 2$ to describe fuzzy functions of two variables; in this case, the list of rules

$$x_1 \,\mathrm{is}\, u_1 \ \wedge \ x_2 \,\mathrm{is}\, u_2 \ \Rightarrow \ y \,\mathrm{is}\, u$$

can be compactly represented by two-dimensional tables

$$\begin{array}{c|c} & \in U_2 \cup \{\mathtt{any}\} \\ \hline \in U_1 \cup \{\mathtt{any}\} & \in U \cup \{\mathtt{any}\} \end{array} \tag{35}$$

whose row headers contain the fuzzy numbers $u_1 \in U_1 \cup \{\mathtt{any}\}$, whose columns headers contain the fuzzy numbers $u_2 \in U_2 \cup \{\mathtt{any}\}$, and whose table entries contain the corresponding result $u \in U \cup \{\mathtt{any}\}$. Default conclusions are specified in the rows and columns labeled `any`; unspecified entries in a table are also taken to be `any`. Usually, a single table (35) is sufficient to define a FAM, but a FAM may also be defined by combining the information from several such tables (typically coming from experts with different preferences).

In many applications, in particular in those to fuzzy control, one is mainly interested in a function $y = f(x)$ consistent with the information in the FAM. We shall call any procedure for extracting such a function from a FAM an **interpretation**. There are a number of different interpretation procedures in the literature, going under the label **defuzzification**, all justified by more or less plausible heuristic arguments.

There are two interesting approaches to the interpretation of a FAM specified in terms of surprise. The first is obtained by application of fuzzy logic. The calculus given in Section 4 implies the following formula for the surprise of y given x:

$$\begin{aligned} s(y|x) &= \sum_{\text{rules}} \frac{s(y|y \,\mathrm{is}\, u)}{1 + s(y|y \,\mathrm{is}\, u) \sum_i s(x_i|x_i \,\mathrm{is}\, u_i)} \\ &= \sum_{\text{rules}} \frac{s_u(y)}{1 + s_u(y) \sum_i s_{u_i}(x_i)}. \end{aligned}$$

Note that surprise functions constructed from the primitives in Table 1 using logical operations are always continuously differentiable. Thus we may use standard software to optimize $s(y|x)$ with respect to y to find the *most plausible value* $f(x) = \hat{y}$ for

y given x. This generally requires for each x the evaluation of $s(y|x)$ and its gradient at a number of points. Thus, while optimal in the stated sense, this interpretation rule is practical only when $f(x)$ is needed only at a few points x, or if there is time enough to do the optimization for each function evaluation needed.

For real time applications when function evaluation must be very fast, one has to be content with a suboptimal interpretation rule. For example, if the statement "y is u" takes the actual form "$y \approx y_u \pm \sigma$" with σ independent of u, we may use the **blending function** defined by

$$f(x) = \sum_{\text{rules}} \frac{y_u}{\varepsilon + \sum_i s_{u_i}(x_i)} \Bigg/ \sum_{\text{rules}} \frac{1}{\varepsilon + \sum_i s_{u_i}(x_i)}, \tag{36}$$

where ε is a tiny number. Here blending refers to the fact that usually several terms from a fuzzy scale (including `any`) apply to the actual value of a variable, and the formula (36) blends the different admitted conclusions in a correctly weighted way, so that it has all the qualitative properties specified in the FAM. Indeed, if some x_i is incompatible with "x_i is u_i" then $s_{u_i}(x_i) = \infty$ and the rule has no influence on $f(x)$. And if the hypothesis of a rule (34) has no surprise at all then $s_{u_i}(x_i) = 0$ for all i and $f(x) \to y_u$ in the limit $\varepsilon \to 0$. Moreover, it is easy to see that, for the surprise function from Table 1, this is a continuously differentiable function. (In contrast, traditional defuzzification rules usually result in nondifferentiable functions.) Similarly, if the statement "y is u" takes the actual form "$y \approx y_u \pm \sigma_u$", we may use the blending function

$$f(x) = \sum_{\text{rules}} \frac{\sigma_u^{-1} f_u(x)}{\varepsilon + \sum_i s_{u_i}(x_i)} \Bigg/ \sum_{\text{rules}} \frac{\sigma_u^{-1}}{\varepsilon + \sum_i s_{u_i}(x_i)}. \tag{37}$$

Note that the blending functions are as easily parallelizable in hardware as are neural networks.

In many cases of practical interest, additional information is available in terms of example pairs (x_l, y_l), where each y_l is considered as an acceptable realization of $f(x_l)$. Typically, these come from illustrative examples and extreme or ambiguous cases provided by one or several experts, or from past performance data. If sufficiently many such pairs are available, one can compute optimal estimates for the tuning parameters occurring in the definition of the elementary surprise functions.

This is done in a **training process** that consists in minimizing the residual sum of squares $\sum_l \|f(x_l) - y_l\|_2^2$ (or a suitably weighted version) with respect to the tuning parameters. The resulting nonlinear least squares problem is tractable since, by our construction of surprise functions, the blending functions (36) and (37) are smooth functions of the tuning parameters. Note that the training process needs to be done only once, and corresponds to training a fuzzy controller before using it online with the optimal tuning parameters obtained by the training process.

Acknowledgment. The author appreciates the careful reading of the manuscript by a referee, whose remarks contributed significantly to a more readable exposition. Thanks also to another referee for providing the historical references mentioned at the end of the introduction.

References

[1] R.G. Cowell, A.P. Dawid, S.L. Lauritzen and D.J. Spiegelhalter, Probabilistic Networks and Expert Systems, Springer, Berlin 1999.

[2] M. Delgado, J. Kacprzyk, J.L. Verdegay and M.A. Vila, Fuzzy Optimization. Recent Advances, Physica-Verlag, 1994.

[3] J. Dombi, A general class of fuzzy operators, the De Morgan class of fuzzy operators and fuzziness measures induced by fuzzy operators, Fuzzy Sets and Systems 8 (1982), 149–163.

[4] D. Driankov, H. Hellendoorn and M. Reinfrank, An Introduction to Fuzzy Control, Springer, Berlin 1993.

[5] D. Dubois, W. Ostasiewicz and H. Prade, Fuzzy sets: history and basic notions, pp. 21–124 in: The Handbook of Fuzzy Sets (D. Dubois and H. Prade, eds.), Vol. 1: Fundamentals of Fuzzy Sets, Kluwer, Dordrecht 2000.

[6] D. Dubois and H. Prade, Fuzzy Sets and Systems: Theory and Applications, Academic Press, New York 1980.

[7] D. Dubois and H. Prade, Possibility Theory, Plenum Press, New York 1988.

[8] D. Dubois and H. Prade, Toll sets and toll logics, pp. 169–177 in: Fuzzy Logic: State of the Art (R. Lowen and M. Roubens, eds.), Kluwer, Dordrecht 1993.

[9] H. Hamacher, Über logische Aggregationen nicht binär expliziter Entscheidungskriterien, Rita G. Fischer Verlag, Frankfurt 1978.

[10] Li-Min Jia and Xi-Di Zhang, Distributed intelligent railway traffic control based on fuzzy decision making, Fuzzy Sets and Systems, 62 (1994), 255–265.

[11] G. Klir and B. Yuan, Fuzzy Sets and Fuzzy Logic. Theory and Application, Prentice Hall, Upper Saddle River, NJ, 1995.

[12] B. Kosko, Neural Networks and Fuzzy Systems, Prentice Hall, Englewood Cliffs, NJ, 1991.

[13] Y-J. Lai and C-L. Hwang, Fuzzy Mathematical Programming. Methods and Applications, Lecture Notes in Economics and Mathematical Systems 394, Springer, Berlin 1992.

[14] W.A. Lodwick, A. Neumaier and F. Newman, Optimization under uncertainity: methods and applications in radiation therapy, Proc. 10th IEEE Int. Conf. Fuzzy Systems 2001, to appear.

[15] R.A. McCain, Fuzzy confidence intervals, Fuzzy Sets and Systems 10 (1983), 281–290.

[16] A. Neumaier, Interval Methods for Systems of Equations, Cambridge Univ. Press, Cambridge 1990.

[17] A. Raggl and W. Slany, A reusable iterative optimization library to solvr combinatorial problems with approximate reasoning, Int. J. Approx. Reasoning 19 (1998), 161-191.

[18] M. Sakawa, Fuzzy Sets and Interactive Multiobjective Optimization, Plenum Press, 1993.

[19] K. Schmucker, Fuzzy Sets, Natural Language Computations, ans Risk Analysis, Computer Science Press, 1984.

[20] G.L.S. Shackle, Decision, Order and Time in Human Affairs, 2nd edition, Cambridge University Press, Cambridge 1961.

[21] P. Shenoy, On Spohn's rule for revision of beliefs, Int. J. Approximate Reasoning 5 (1991), 149–181.

[22] W. Slany, Scheduling as a Fuzzy Multiple Criteria Optimization Problem, Fuzzy Sets and Systems 78 (1996), 197–222.

[23] W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states, pp. 105–134 in: Causation in Decision, Belief Change and Statistics (W. Harper and B. Skyrms, eds.), Kluwer, Dordrecht 1988.

[24] W. Spohn, A general non-probabilistic theory of inductive reasoning, pp. 149–158 in: Uncertainty in Artificial Intelligence 4 (R.D. Shachter, T.S. Levitt, L.N. Kanal and J.F. Lemmer, eds.), North-Holland, Amsterdam 1990.

[25] H.-J. Zimmermann, Fuzzy Sets, Decision Making, and Expert Systems, Kluwer, Boston 1987.

[26] H.-J. Zimmermann, Fuzzy Set Theory – and Its Applications, 2nd ed., Kluwer, Dordrecht 1991.

[27] H.-J. Zimmermann, The application of fuzzy set theory to mathematical programming, Information and Control 36 (1985), 29–58.