1

NESTED INTERVALS AND SETS: CONCEPTS, RELATIONS TO FUZZY SETS, AND APPLICATIONS

Hung T. Nguyen* and Vladik Kreinovich**

** Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM 88003-8001, email hungueyn@nmsu.edu*

*** Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, email vladik@cs.utep.edu*

ABSTRACT

In data processing, we often encounter the following problem: Suppose that we have processed the measurement results $\tilde{x}_1, ..., \tilde{x}_n$, and, from this processing, have obtained an estimate $\tilde{y} = f(\tilde{x}_1, ..., \tilde{x}_n)$ for a quantity $y = f(x_1, ..., x_n)$; we know the intervals $\mathbf{x}_i$ of possible values of x_i, and we want to find the interval $\mathbf{y}$ of possible values of y. Interval computations are one of the main techniques for solving this problem.

In some cases, for each i, in addition to the *guaranteed* interval $\mathbf{x}_i$ of possible values, we have a smaller interval that an expert believes to contain x_i. There may be several such *nested* intervals. In these cases, in addition to the guaranteed interval $\mathbf{y}$, it is desirable to know the possible intervals of y that correspond to the opinions of different experts.

Techniques of such *nested interval computations* and real-life applications of these techniques are described in this paper.

*This work was partially carried out while Hung T. Nguyen was on sabbatical leave at the University of Southern California, Los Angeles, Spring 1995

1 INTRODUCTION

1.1 Estimating Errors of Indirect Measurements: One of the Main Problems in Data Processing

One of the important problems in data processing is to estimate the error of an indirect measurement. Namely, we want to know a value of the physical quantity y that is difficult or impossible to measure directly. So, to estimate the values of y, we measure other quantities $x_1, ..., x_n$ that are related to y, and reconstruct y from the results $\tilde{x}_i$ of these measurements. This procedure is called an *indirect measurement* .

It is usually known how to reconstruct y from x_i. In other words, we usually know an algorithm f that, if applied to the actual values of x_i, would result in the actual value of $y = f(x_1, ..., x_n)$. Instead of the actual values x_i, we only know the measurement results $\tilde{x}_i$ that (due to measurement errors) may be slightly different from the actual values. As a result, when we apply the algorithm f to these measurement result, the resulting estimate $\tilde{y} = f(\tilde{x}_1, ..., \tilde{x}_n)$ for y is, generally speaking, different from the actual values of y.

If we want to know what the actual value of y can be, we must know how different y can be from this estimate $\tilde{y}$. In other words, we need to estimate the error $\Delta y = \tilde{y} - y$ of this indirect measurement.

This error Δy comes from the errors $\Delta x_i = \tilde{x}_i - x_i$ of measuring x_i. To measure x_i, we use regular measuring instruments. For a measuring instrument to make sense, the manufacturer must provide it with some guaranteed error bounds Δ_i. (Otherwise, whatever value $\tilde{x}_i$ we measure, the actual value x_i can be arbitrarily far from $\tilde{x}_i$, i.e., no information about x_i would come from the measurement itself.) If we know the bound Δ_i, and the result of the measurement is $\tilde{x}_i$, then the actual value must belong to the interval $[\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$.

In some cases, the manufacturer also provides the user with the *probabilities* of different values of error Δx_i, but in many real-life cases, the bounds Δ_i are the only information about the errors that we have. In these cases, we have the following problem:

We know:

- *the intervals* $\mathbf{x}_i = [\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$ *that contain the actual values of* x_i*;*
- *an algorithm* f *that transforms* n *values* $x_1, ..., x_n$ *into a real number.*

We must compute: The set $\mathbf{y}$ *of possible values of* y*:* $\mathbf{y} = f(\mathbf{x}_1, ..., \mathbf{x}_n) = \{f(x_1, ..., x_n) \,|\, x_1 \in \mathbf{x}_1, ..., x_n \in \mathbf{x}_n\}$.

Comments.

1. This is the *main problem* to be considered in the present paper. In addition to the problem, we will also consider other problems, such as identification, optimization, design, and control. These problems will be briefly outlined in Section 3.

2. In some cases, the manufacturer can guarantee that some values of error are impossible. In this case, the set $\mathbf{x}_i$ of possible actual values can be a subset of the interval $[\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$.

3. A measuring instrument can measure several different quantities $y_1, ..., y_k$ at a time. In this case, in addition to the information about the possible errors of each measurement, the manufacturer can guarantee that certain combinations of errors are impossible: e.g., it can happen that Δy_1 attains its largest possible value Δ, and it can happen that Δy_2 attains its largest possible value Δ, but they can never attain these extreme values at the same time, because of the restriction that $\Delta y_1^2 + \Delta y_2^2 \le \Delta^2$. (This situation happens, e.g., if we measure geographical coordinates of a point.) Such information can be described by a set $\mathbf{y} \subset R^k$ of all the tuples that the manufacturer believes can be possible values of errors $(\Delta y_1, ..., \Delta y_k)$.

1.2 Input Intervals are Often Overestimated

Input intervals come from the manufacturer's guarantee. The manufacturer does not want to risk his/her reputation, and to be liable to lawsuits. So, when in doubt, the manufacturer prefers to overestimate the error bounds Δ_i. The actual designers and testers of the measuring instrument, when you ask their opinion, usually assign much better quality to their instruments. They can say something like "We can only *prove* that the error is ≤ 0.1, but all our experiments have shown that this error is ≤ 0.05."

Cautious overestimation is a reasonable strategy when the error may be life-threatening: e.g., in measuring the reactor characteristics of the nuclear power stations, or in computing the trajectory of the spaceship. But this caution makes no sense if we are, e.g., estimating the amount of oil in a given area. In such cases, instead of an absolutely accurate and absolutely guaranteed estimate like 100 ± 50 mln tons, it would be nice to know that with high reliability, we are able to extract from the same measurement results a much narrower interval (e.g., 90 ± 10).

1.3 In Addition to the Main (Guaranteed) Interval, it is Reasonable to Consider its Subintervals

In these situations (when nothing life threatening can come from an underestimation of a measurement error), it is thus reasonable, in addition to the intervals that are *guaranteed* by the manufacturer, to consider *subintervals* that the manufacturer cannot, strictly speaking, guarantee, but which the experts (designers and/or producers of the manufacturing instrument) claim to be true.

In other words, inside each interval $\mathbf{x}_i$, there is another interval $\mathbf{x}_i(c) \subseteq \mathbf{x}_i$.

Example. Suppose that we have a measuring instrument with guaranteed accuracy $\Delta = 0.1$ and actual accuracy 0.06. (We give this value to illustrate the idea; in real life, this value is not exactly known neither to the manufacturer of the measuring instrument, nor to the user.) Suppose also that, in addition to the guaranteed error bound 0.1, we have an expert's estimate that the error bound cannot exceed 0.08. Suppose that we have performed the measurement, and the measurement result is $\tilde{x} = 1.1$. In this case, the guaranteed estimate $\mathbf{x} = [\tilde{x} - \Delta, \tilde{x} + \Delta] = [1.0, 1.2]$. The actual (unknown) set of possible values is $X_{\rm act} = [1.1 - 0.06, 1.1 + 0.06] = [1.04, 1.16]$. The set of possible values according to the expert is equal to $\mathbf{x}(c) = [1.1 - 0.08, 1.1 + 0.08] = [1.02, 1.18]$.

Given these additional n intervals, we would like, in addition to the *guaranteed* interval $\mathbf{y} = f(\mathbf{x}_1, .., \mathbf{x}_n)$, to compute the smaller (*claimed*) interval for y: $\mathbf{y}(c) = f(\mathbf{x}_1(c), ..., \mathbf{x}_n(c))$.

A first time reader may ask himself/herself: "Why consider subintervals that are not guaranteed by the manufacturer? Are there other motivations for using nested sets other than experts suspecting the manufacturer's error bounds are

overestimates?" To explain the practical usefulness of nested sets, let us give a couple of real-life examples where such subintervals are useful (in Section 4, we will show how nested interval computations help to solve the corresponding real-life problems).

Case study: seismic engineering (Dong et al [13]). Earthquake are extremely difficult to predict. Because of that, if we only use the equations and the measured data x_i, we get very wide intervals of possible magnitudes and frequencies y, intervals that are known to be much wider than the interval of actual values. So, if we use these wide intervals $[y^-, y^+]$ to design a building that is guaranteed to withstand a typical earthquake, these buildings will be unnecessarily (and often unrealistically) expensive. To make the requirements more realistic, in addition to guaranteed measurement results, *expert estimates* are normally used that lead to narrower intervals.

In the second example, measurements are so bad that we have to rely mainly on expert estimates:

Case study: AIDS epidemics (Berleant et al [6]). Equations that describe how AIDS is transferred are known. To solve these equations, we must know the initial conditions and the the numerical values of the parameters (such as the death rate among AIDS patients, etc). The majority of these parameters can be determined more or less accurately from the known statistical data. However, there is one important parameter about which it is very difficult to get a reliable statistical estimate: the number of intravenous drug abusers. So, in the absence of a reliable statistical estimate, we may use the expert estimates. Different experts can give different intervals of possible values, so, we have nested intervals for describing the expert knowledge.

1.4 There Can Be Several Subintervals

In the previous subsections, we assumed that there are two intervals that correspond to two degrees of confidence:

- an interval guaranteed by the manufacturer; this interval corresponds to an *absolute* confidence (100% guarantee); and
- its subinterval, that contains only the values that the expert (e.g., designer) believes to be possible.

This is, however, an idealized situation. Different experts can have different ideas of what values are possible and what values are not. More optimistic experts will choose smaller intervals, while more cautious experts will prefer larger ones. As a result, for each variable x, we may have several different intervals of possible values $\mathbf{x}(\alpha)$ that depend on an expert α (i.e., we have a *nested* interval). For each α, we end up with a different interval $\mathbf{y}(\alpha)$. So, we get a nested interval for the result. In this paper, we describe the methods of computing nested intervals, and applications of these computations.

At first, we consider the case when experts can be ordered in a sequence form the most cautious one to the most risky one. Then, we will consider the general case when some experts may be incompatible.

2 NESTED INTERVALS AND NESTED SETS

2.1 Nested Sets: Motivation of the Definition

For simplicity, let us first consider the case when the experts can be ordered in a sequence from the most cautious to the most risky. (In Section 2.5, we will mention what happens if we do not make this assumption.)

Let us assume that we have the results $\tilde{x}$ of one or several measurements. We want to know what information about the actual values x we can deduce from $\tilde{x}$. In other words, we want to know what values of x could possibly lead to these measurement results. In more mathematical terms, we want to know the set X_{act} of all values of x that could lead to these measurement results.

We do not know this set precisely. Instead, we know the following: first, from the manufacturer's guarantee, we know some set that definitely contains X_{act} (i.e., a set with the property that every possible value x definitely belongs to it). For a single measurement with an error bound Δ, such a set is an interval $[\tilde{x} - \Delta, \tilde{x} + \Delta]$. In general, we will describe this set by $X(\alpha_0)$. (The reason for choosing this notation will be given later.)

We believe that $X(\alpha_0)$ is possibly an *overestimate* for X_{act}, so that the actual set of possible value is probably smaller than $X(\alpha_0)$. This belief is shared by several experts, who propose their opinions on the set of possible values. Let us

denote the number of experts by d. We have assumed that the experts can be ordered in the order of their cautiousness, from the most cautious one (# 1 in this ordering), to the second cautious (# 2 in this ordering), ..., to the riskiest expert (who in this ordering gets # d). Let us denote the set of all values x that the $j-$th expert believes to be possible by $X(\alpha_j)$. Then, the larger j, the riskier the expert's statement becomes, and therefore, the fewer elements are called possible. In mathematical terms, this means that we have the following relationship between these sets: $X(\alpha_0) \supseteq X(\alpha_1) \supseteq X(\alpha_2) \supseteq ... \supseteq X(\alpha_d)$.

If an element x belongs to the set $X(\alpha_j)$, this means that not only the $j-$th expert believes this element to be possible, but actually j experts (1, 2, ..., j) all agree that x is a possible value. So, elements from the set $X(\alpha_j)$ can be described as elements about which at least j out of d experts agree that x is possible. So, if we would estimate the degree of possibility of an element x by polling experts, elements from $X(\alpha_j)$ would be those for which at least j experts voted. The more experts vote for x to be possible, the larger is the expert's community *degree of belief* that x is possible. In these terms, we can say that the sets $X(\alpha_j)$ describe different *degrees of belief* in the possibility of x.

In addition to an *overestimate* $X(\alpha_0)$, we may also know that some values x are *definitely possible*. For example, we may estimate Δ by adding up the upper bounds for component errors, but we may also be sure that a certain error component is definitely there (e.g., thermal noise), so we may know a lower bound for the largest possible error. If we know a set D whose values are definitely possible, then, of course, experts' estimates $X(\alpha_j)$ must contain this set D. It is therefore smaller than all $X(\alpha_j)$, i.e., we have $X(\alpha_0) \supseteq X(\alpha_1) \supseteq ... \supseteq X(\alpha_d) \supseteq D$. This ordering prompts a natural notation $X(\alpha_{d+1})$ for this set D.

Example. Let us illustrate these notions on the same example as we used before to explain subintervals. Suppose that we have a measuring instrument with guaranteed accuracy $\Delta = 0.1$ and (unknown) actual accuracy 0.06. The only lower bound for the error bound is 0. Suppose also that we have two experts: a cautious one that estimates the error bound as 0.08, and a risky one that estimates it as 0.05. Here, $d = 2$. Suppose that we have performed the measurement, and the measurement result is $\tilde{x} = 1.1$. In this case, the guaranteed estimate $X(\alpha_0) = [\tilde{x} - \Delta, \tilde{x} + \Delta] = [1.0, 1.2]$. The actual (unknown) set of possible values is $X_{\rm act} = [1.1 - 0.06, 1.1 + 0.06] = [1.04, 1.16]$. The sets of possible values according to the two experts are: $X(\alpha_1) = [1.1 - 0.08, 1.1 + 0.08] = [1.02, 1.18]$ and $X(\alpha_2) = [1.1 - 0.05, 1.1 + 0.05] = [1.05, 1.15]$. Finally,

the set $D = X(\alpha_3)$ that contains elements that are guaranteed to be possible consists of a single element $X(\alpha_3) = \{1.1\}$.

Suppose now that we make a second measurement of the same quantity, and get a new result 0.99. This means that we can guarantee that

$$x \in [0.99 - 0.1, 0.99 + 0.1] = [0.89, 1.09].$$

Combining this information with the guaranteed information $x \in [1.0, 1.2]$, we can conclude that $x \in [1.0, 1.2] \cap [0.89, 1.09] = [1.0, 1.09]$, so, $[1.0, 1.09]$ is our new set $X(\alpha_0)$. For the first expert, the same procedure leads to $X(\alpha_1) = [1.02, 1.18] \cap [0.99 - 0.08, 0.99 + 0.08] = [1.02, 1.07]$. For the new set D, we have the following: if we only had the result of the first measurement, then we could guarantee that the value 1.1 is possible. However, now that we have the additional information, the set of possible values shrinks, and there is no guarantee that 1.1 is possible anymore. (Actually, it is not, because the error is bounded by 0.06, so if the actual value was 1.1, we could not have 0.99 as a measurement result.) Mathematically speaking, from the first measurement, we get $D = \{1.1\}$, from the second, we get $D = \{0.99\}$. In this case, the intersection $\{1.1\} \cap \{0.99\}$ is empty, and therefore, there is no value that we can guarantee to be possible. So, in this case, $X(\alpha_3)$ is empty. Similarly, $X(\alpha_2) = [1.05, 1.14] \cap [0.99 - 0.05, 0.99 + 0.05] = \emptyset$ (which means that the opinion of the second expert turned out to be incorrect).

Comment. Different values of α may sometimes describe the opinion not of different experts, but of one and the same expert: if we ask for an expert's most reliable opinion, the expert may give a much broader interval, but if we ask the expert for what she really thinks (thus allowing her to err), the same expert may give us a narrower interval.

Let us describe nested sets and nested intervals formally.

2.2 Definition and the Formulation of the Main Problem

Definition 1.

- *Let an integer d be fixed. This number will be called the number of experts.*
- *Let a finite linearly ordered set $A = \{\alpha_0 < \alpha_1 < \alpha_2 < ... < \alpha_d < \alpha_{d+1}\}$ be fixed. The elements of this set will be called degrees of belief.*

- *Let a set U be given. This set will be called the Universal set, or, the Universe of discourse.*
- *By a piece of information (or, a nested set), we mean a non-increasing mapping X from A to the set 2^U of all subsets of U, i.e., a tuple $(X(\alpha_0), X(\alpha_1), ..., X(\alpha_{d+1}))$, for which $X(\alpha_{d+1}) \subseteq X(\alpha_d) \subseteq ... \subseteq X(\alpha_1) \subseteq X(\alpha_0) \subseteq U$.*
- *Let u be an element of U, and $\alpha \in A$ be a degree of belief. We say that u belongs to the nested set X with a degree of belief $\geq \alpha$ if $u \in X(\alpha)$.*

Historical comment. The notion of a nested set was first introduced in Gentilhomme [20] under the French name "ensemble flou". It was further analyzed by Negoita et al [51, 52].

Two Problems: Finding the Input Nested Sets and Computing the Nested Set for the Result

Our ultimate goal is to describe the nested set Y that characterizes y. To do that, we must find the nested sets X_i for all input variables x_i, and then transform these input nested sets into a nested set Y for y. Before we describe the algorithms, we need to describe these problems in mathematical terms.

Finding the Input Nested Set

Finding the nested set X that describes the input variable x may not be a problem if we have a single measurement result. However, in some cases, we measure the same quantity several times, and thus obtain several measurement results that lead to several pieces of information (nested sets) X, X', ... So, we must combine them to get a nested set that describes all the information about the input variable x.

If we will be able to combine two nested sets, then, we will be able to combine any number of them X, X', X'', ...:

- we start by combining the first two nested sets (X and X');
- then, we combine their combination with the third nested set (X''), etc.

Therefore, in the following text, we will only describe how to combine two nested sets X and X'. This combination can be easily done: Indeed, let us consider a degree of belief α_j. If $1 \le j \le d$, then this degree of belief describes the opinion of the $j-$th expert. For this expert, according to the first piece of information, the actual value x belongs to the set $X(\alpha_j)$, and according to the second piece of information, x belongs to $X'(\alpha_j)$. As a result, for this expert, the set of possible values of x coincides with $X(\alpha_j) \cap X'(\alpha_j)$. A similar formula can be derived for the case when α describes the guaranteed set of possible values. So, we arrive at the following definitions:

Definition 2. *By an intersection $X \cap X'$ of nested sets X and X', we mean a nested set with $(X \cap X')(\alpha) = X(\alpha) \cap X'(\alpha)$.*

In some cases, the reason we perform several measurements is, e.g., that one of them may err, so that only one of them is reliable. For two measurements, this means that after we get pieces of information X and X', we can only be sure that the actual value x belongs to one of the corresponding sets. For an expert α, this means that set of possible values of x is a *union* $X(\alpha) \cup X'(\alpha)$. So, we can define a union of nested sets:

Definition 3. *By a union $X \cup X'$ of nested sets X and X', we mean a nested set with $(X \cup X')(\alpha) = X(\alpha) \cup X'(\alpha)$.*

There are also situations when another set operation naturally arises: complementation of the two sets. These situations are slightly more complicated. As an example, let us assume that the measurement error is caused by noise. We know that this noise is always present, but we do not know whether this noise is positive or negative. In this case, the set of possible values of the noise is a union of two intervals: one that corresponds to the positive values, and one that corresponds to the negative values. For simplicity, let us assume that this set of symmetric, i.e., it is of the form $[-\Delta, -\delta] \cup [\delta, \Delta]$ for some $0 < \delta < \Delta$. This set can be represented as a difference between the two intervals: $[-\Delta, \Delta] \setminus [-\delta, \delta]$. For each of these intervals, we can have more cautious and more risky estimates in the above sense. A more cautious estimate for Δ means a larger Δ; similarly, a more cautious (in the sense of Definition 1) estimate for δ also means a larger δ. However, the most cautious estimate for the difference corresponds to the case when we take the *largest* possible Δ and the *smallest* possible δ. In other words, to get the most cautious estimate of the difference $[-\Delta, \Delta] \setminus [-\delta, \delta]$, we subtract the *riskiest* estimate for $[-\delta, \delta]$ from the *most cautious* estimate for $[-\Delta, \Delta]$. So, to describe negative information (for set complement), the expert uses not his/her own degree of belief α, but some other degree of belief $\alpha^{\perp}$: the

most cautious α uses the riskiest $\alpha^\perp$, the more cautious α uses the riskier $\alpha^\perp$, and the riskiest α uses the most cautions $\alpha^\perp$.

This description works for the values α_j that correspond to different experts. A similar idea works for the values $j = 0$ and $j = d+1$ that describe the upper and lower bounds for the set of possible values: Indeed, we can view the sets $X^\perp(\alpha_j)$ as a description of the complement $X^\perp_{act}$ to the (unknown) set $X_{\rm act}$. The only information that is guaranteed for $X_{\rm act}$ is that $X(\alpha_{d+1}) \subseteq X_{\rm act} \subseteq X(\alpha_0)$. From these inclusions, we can conclude that $X^\perp(\alpha_0) \subseteq X^\perp_{act} \subseteq X^\perp(\alpha_{d+1})$. So, $X^\perp(\alpha_0)$ is a guaranteed lower bound for the desired (unknown set) $X^\perp_{act}$, and $X^\perp(\alpha_{d+1})$ is a guaranteed upper bound. So, we can take $X^\perp(\alpha_0)$ as $(X^\perp)(\alpha_{d+1})$, and $X^\perp(\alpha_{d+1})$ as $(X^\perp)(\alpha_0)$. In other words, for $\alpha = \alpha_0$ and $\alpha = \alpha_{d+1}$, we define $(X^\perp)(\alpha) = X^\perp(\alpha^\perp)$, where $\alpha_0^\perp = \alpha_{d+1}$ and $\alpha_{d+1}^\perp = \alpha_0$. Let us summarize this description in mathematical terms.

Definition 4. *Let $\alpha \to \alpha^\perp$ denote a strictly decreasing function $A \to A$ for which $\alpha_0^\perp = \alpha_{d+1}$ and $\alpha_{d+1}^\perp = \alpha_0$. By a complement $X^\perp$ to a nested set X, we mean a nested set $(X^\perp)(\alpha) = X^\perp(\alpha^\perp)$ (where $A^\perp$ means a complement $U \setminus A$). By a difference $X \setminus X'$ between nested sets X and X', we mean a nested set $X \cup (X')^\perp$.*

The desired functions $\alpha \to \alpha^\perp$ can be easily described:

PROPOSITION 1. *If $\alpha \to \alpha^\perp$ is a strictly decreasing function $A \to A$ for which $\alpha_0^\perp = \alpha_{d+1}$ and $\alpha_{d+1}^\perp = \alpha_0$, then $\alpha_j^\perp = \alpha_{d+1-j}$.*

(For the reader's convenience, all proofs are placed in a special Appendix).

Thus defined complement operations have some properties of the set-theoretic complement: e.g., De Morgan laws $(X \cup X')^\perp = X^\perp \cap (X')^\perp$ and $(X \cap X')^\perp = X^\perp \cup (X')^\perp$ are easy to prove. However, some properties of this complement operation may sound counterintuitive: for example, it is possible that $X \setminus X \neq \emptyset$ and $X \cup X^\perp \neq U$. This contradiction with traditional set theoretic intuition can be easily explained if we use Definition 4: e.g., $X \setminus X$ does not mean that we subtract a set X from itself. To get $(X \setminus X)(\alpha)$, we subtract from $X(\alpha)$ the set $X(\alpha^\perp)$ that corresponds to a *different* degree of belief, and can thus be *smaller* than $X(\alpha)$ (if $\alpha^\perp > \alpha$).

Computing the Nested Set for y

The result of processing nested sets can be defined expert-by-expert:

Definition 5. *Let $X_1, ..., X_n$ be nested sets with a common set A, and let $f : U \times ... \times U \to U$ be a function. Then, we define $f(X_1, ...X_n)$ as a nested set Y with $Y(\alpha) = f(X_1(\alpha), ..., X_n(\alpha))$.*

The *main problem* of nested set processing can be thus formulated as follows: *given* the pieces of information $X_1, ..., X_n$ that describe inputs, *compute* the nested set Y that describes the output.

2.3 Preliminary Step: Checking Whether Expert Knowledge is Consistent

An expert may be wrong. In some cases, it is possible to check that: namely, if we know that the measured quantities x_i must satisfy some constraints (equalities or inequalities). If for $x_i \in X_i(\alpha)$, these constraints cannot be satisfied, this means that an expert α is wrong. For example, assume that we are measuring three currents, and we know that due to Kirkhoff's law, $x_3 = x_1 + x_2$. Suppose that as a result of measuring these currents, we got the values $\tilde{x}_1 = 1.0$, $\tilde{x}_2 = 2.0$, and $\tilde{x}_3 = 2.9$. If the measurement accuracy is 0.1, then the constraint can still be satisfied: e.g., for $x_1 = 1$, $x_2 = 2$, and $x_3 = 3$. Assume, however that an expert α believes that the actual accuracy of the measuring instrument is 0.01. Then, for this expert, $X_1(\alpha) = [0.99, 1.01]$, $X_2(\alpha) = [1.99, 2.01]$, and $X_3(\alpha) = [2.89, 2.91]$. In this case, $X_1(\alpha) + X_2(\alpha) = [2.97, 3.03]$, so $X_3(\alpha) \cap X_1(\alpha) + X_2(\alpha) = \phi$, and these conditions are inconsistent. So, this expert is wrong. In Mitra et al [44], an algorithm is proposed that checks consistency for the case when x_i corresponds to measuring time.

2.4 Methods of Solving the Main Problem

A Natural Approach: Computing $Y(\alpha)$ Set-by-Set (Expert-by-Expert). Advantages of the Natural Approach

A natural way of computing Y is to compute $Y(\alpha_j)$ for $i = 0, 1, ..., d + 1$ with traditional interval techniques (if the sets $X_i(\alpha)$ are intervals), or generalized interval techniques (if the sets $X_i(\alpha)$ are more complicated than intervals). This method is described, e.g., in Gerasimov et al [21]. This approach has two advantages:

- it uses well-known techniques for which good software tools have been developed; and
- it is easily parallelizable: if we have $d+2$ processors $P_0, ..., P_{d+1}$, then we can let each of them compute $Y(\alpha_j)$ for the corresponding j; if the number of processors p is smaller than $d+2$, then we can divide A into p groups, and ask $m-$th processor to compute the values $Y(\alpha_j)$ for all j from the $m-$th group.

An Alternative Approach

To describe how nested sets are processed, we must first describe how they are stored in the computer. The natural approach assumes that we store each nested set as a sequence of sets $X(\alpha_0), ..., X(\alpha_{d+1})$. When all the sets $X(\alpha_j)$ are intervals, this is a reasonable idea: each interval is stored as two numbers (left- and right endpoints), so, to store $d+2$ intervals, we store $2(d+2)$ numbers. To describe a nested interval, we need exactly that many numbers, so, there is no redundancy in this storage scheme.

However, if the sets $X(\alpha_j)$ are more complicated than intervals, then this expert-by-expert storage scheme is not the best one. Indeed, how do we store a set S in the computer? If the Universe of discourse U is finite, then we can represent a set S by storing, for each $u \in U$, the value 1 or 0 ("true" and "false") depending on whether u belongs to the set S or not. (This is, e.g., how set type in Pascal is represented.) In other words, a set S is stored as the values of its *characteristic function* $\chi_S(u)$.

So, to store $d+2$ sets $X(\alpha_j)$, we must, for every u, store $d+2$ truth values $\chi_{X(\alpha_j)}(u)$. So, for each $u \in U$, we need d bits to store all the information about this u. But this is wasting storage space. Indeed, we know that the sequence $X(\alpha_j)$ is monotonic: $X(\alpha_j) \subseteq X(\alpha_k)$ if $j > k$. Therefore, if for some element j, we know that $u \in X(\alpha_j)$, then we automatically know that $u \in X(\alpha_{j-1})$, ..., $u \in X(\alpha_0)$; we do not need to waste storage space to store these values. Similarly, if we know that $u \notin X(\alpha_j)$, then we can automatically deduce that $u \notin X(\alpha_{j+1})$, ..., $u \notin X(\alpha_{d+1})$.

In view of this monotonicity, for every $u \in U$, there are only $d+3$ possible situations:

- $u \notin X(\alpha_j)$ for all j;
- $u \in X(\alpha_j)$ and $u \notin X(\alpha_{j+1})$ for some $j = 0, ..., d$;
- $u \in X(\alpha_j)$ for all j.

Since there are $d+3$ cases, we can enumerate them by numbers from 0 to $d+2$, and describe the case in the computer by its number. To store a number from 0 to $d+2$ in its binary form, we need $\approx \log_2(d+2)$ bits. For large d, we have $\log_2(d+2) \ll d$ bits. So, this is a much better storage scheme.

Hence, a better way to *describe* (*store*) a nested set is for each $u \in U$, to store the number of the case that describes this particular u.

Comment. If a set U is infinite, then we cannot store all the values $\chi_S(u)$ of the characteristic function. Instead, we have a *program* that, given u, returns the value of $\chi_S(u)$. For the new storage scheme, we will similarly need a function that would input u and return the corresponding case number.

This alternative representation of nested sets not only saves computer *memory*, but it can also lead to *faster algorithms*. Indeed, in modern computers, the actual time to perform arithmetic operations with numbers that are already in the processor is negligible in comparison with the time that is necessary to fetch these numbers from the memory to the processor and move the results back to the memory. Data processing algorithms are often reasonably simple, so that few arithmetic operations are necessary; therefore, the major part of the computation time for these algorithms is taken by this fetching. For such data processing algorithms, the fewer memory cells we need to store the data, the faster the resulting algorithms. Therefore, for the alternative representation, the algorithms can be faster.

We should warn the reader that although we practically *proved* that the alternative representation leads (asymptotically) to better storage, the reduction in computation time is justified only by, so to say, *physical* (less formal) arguments. These arguments seem to infer that in some cases, the alternative representation leads to faster computations, but, on the other hand, in other cases (e.g., if the data processing algorithm is complicated enough) expert-by-expert computations may turn out to be faster.

There is also a practical argument that the alternative representation is often faster: In the following sections, we will describe how to define and process nested sets stored in this manner. It turns out that this alternative way of

storing nested sets corresponds to *fuzzy sets* and operations with these sets. So, numerous successful applications of fuzzy techniques are a justification that this alternative representation is indeed often faster.

Mathematical comment. From the mathematical viewpoint, it would be nice to have a clear realistic example in which the computation time is estimated for both approaches, and in which the alternative leads to a faster processing algorithm; unfortunately, there is no such example in the literature. We feel that, no matter how successful the applications are, from the mathematical viewpoint, the absence of such an example is a major drawback of this approach. We hope therefore that such an example will eventually appear.

Comment. According to the alternative representation, to *describe* (*store*) a nested set, for each $u \in U$, we store the number of the case that describes this particular u. A natural way to enumerate these cases is to use the value j (or α_j) which is the largest α for which $u \in X(\alpha_j)$. This definition also covers the last case (when $u \in X(\alpha_j)$ for all j). However, it does not cover the first case, when no such α exists. For sets of real numbers, the supremum of the empty set is defined as $-\infty$, i.e., as a special value that is added to a real line, and that has the property that it is smaller than everything on the real line. We can make a similar addition to our ordered set A, and arrive at the following definitions:

Mathematical Formulation of the New Approach

Definition 6. *For every ordered set A, let us define A^* as $A \cup \{\alpha_{-1}\}$, where $\alpha_{-1} < \alpha$ for every $\alpha \in A$. For every set $B \subseteq A$, we define* $\sup B$ *as the largest element of B, and* $\sup \emptyset$ *as α_{-1}. For every nested set X, we define $\mu_X(u)$ as* $\sup\{\alpha | u \in X(\alpha)\}$. *This function μ_X is called a membership function of the nested set X.*

In other words, $\mu_X(u)$ is the largest value α for which the degree of belief that $u \in X$ is $\geq \alpha$. The largest value with this property is clearly the degree of belief itself. So, the values $\mu_X(u)$ is called the *degree of belief* that $u \in X$.

In other words, we represent a nested set by a function $\mu : U \to A^*$. That this description is not redundant can be seen from the fact that such functions are in 1-1-correspondence with nested sets:

PROPOSITION 2. *Every function $\mu : U \to A^*$ is a membership function of some nested set.*

To use this type of storage, we must be able to do the following:

- First, we must reformulate all operations with nested sets in terms of membership functions.
- Second, we must be able to reconstruct the nested set from its membership function.

PROPOSITION 3.
$\mu_{X\cap X'}(u) = \min(\mu_X(u), \mu_{X'}(u))$;
$\mu_{X\cup X'}(u) = \max(\mu_X(u), \mu_{X'}(u))$,
$\mu_{X^\perp}(u) = f_\neg(\mu_X(u))$, *where* $f_\neg(\alpha_j) = \alpha_{d-j}$.

PROPOSITION 4. (Nguyen [53])

$$\mu_{f(X_1,...,X_n)}(u) = \max_{u_1,...,u_n : f(u_1,...,u_n)=u} \min(\mu_{X_1}(u_1), ..., \mu_{X_n}(u_n)).$$

PROPOSITION 5. *For every nested set* X, $X(\alpha) = \{u|\mu_X(u) \geq \alpha\}$.

Historical Comment. In view of these propositions, a nested set can be defined as a function $\mu : U \to A^*$. For several nested sets, the union and intersection can be defined as in Proposition 3, and the result of applying a function to nested sets – as in Proposition 4. These definitions have been originally proposed by L. Zadeh (on different grounds) in [78], with the only difference that what we called *nested set* Zadeh called a *fuzzy set*. The set $\{u|\mu_X(u) \geq \alpha\}$ is called an $\alpha-$*level set*, or $\alpha-$*cut* of the fuzzy set X. Formula from Proposition 4 is called *extension principle*. It was proposed by Zadeh in his pioneer paper [78] (see also [79, 80, 16, 68, 33]).

Thus, we get an *explanation* of Zadeh's formulas. Different parts of this explanation appeared in Gentilhomme [20], Tanaka et al [63, 64], Negoita et al [51, 52], Orlovsky [56, 57], Nguyen [53], Dubois et al [16, 17], Kaufmann et al [29], Artbauer [1], Gu et al [27], Goodman et al [24, 25], Klement [31], Lai [36], Tamura et al [62], Uehara et al [66], Bojadziev et al [11], Kruse et al [35], Grabisch et al [26], Klir et al [33], Mantaras [42].

Comments.

- Thus, nested sets provide a new justification of the formulas of fuzzy set theory, formulas that have been originally proposed on a more intuitive

basis, without a convincing mathematical justification. We hope that this new justification will be more convincing, at least for specialists in interval computations.

- Instead of describing a function with finitely many values, it is often more convenient to approximate it by a continuous function. In other words, we can generalize it to the case of *infinite* set A. Generalization of Proposition 4 to infinite A and related problems are described in Nguyen [53] and Fuller et al [18].

How to Store Degrees of Belief in the Computer

Different degrees of belief can be interpreted as ranging from "absolutely false" (α_{-1}) to "absolutely true (α_{d+1}). In the computer, "false" is usually represented by 0, and "true" by 1. So, it is reasonable to represent intermediate degrees of belief as real numbers from 0 to 1 in such a way that α_{-1} corresponds to 0, and α_{d+1} to 1. The most natural arrangement is when the resulting values are equally spaced, i.e., if $\alpha_{-1} = 0$, $\alpha_0 = 1/(d+2)$, ..., $\alpha_i = (i+1)/(d+2)$, ..., $\alpha_{d+1} = 1$. These values can be interpreted as follows: Indeed, before we consider our d experts with their additional knowledge, we have only two options: either we can be pessimistic, and assume that every value that can potentially be possible is actually possible (in this case, our set of possible values is $X(\alpha_0)$), or we can be extremely optimistic and assume that only those values about which we are 100% sure that they are possible are actually possible, and none else (in this case, our set of possible values is $X(\alpha_{d+1})$). After we add d experts, we now have $d+2$ possible viewpoints: a pessimistic one, an optimistic one, and d viewpoints of different experts.

It is easy to see that for every value $u \in U$ and for every nested set X, $\mu_X(u)$ is equal to the ratio of the two numbers:

- the numerator $i+1$ is the number of viewpoints according to which u is a possible value of x, and
- the denominator is the total number $d+2$ of possible viewpoints.

Indeed, if $\mu_X(u) = \alpha_i = (i+1)/(d+2)$, $1 \le i \le d$, this means that the $i-$th is the most optimistic expert who believes that u is possible. Therefore, since we have assumed that the experts are ordered according to their cautiousness, i experts $1, ..., i$ believe that u is possible, and also, in this case, a pessimist

would believe that u is possible. So, u is possible according to $i+1$ out of $d+2$ viewpoints. Similarly, this equality can be checked for $i=-1$, $i=0$, and $i=d+1$.

This equality leads to the possibility to determine $\mu_X(u)$ by *polling*. To do the polling, we need to get experts who represent different points of view: both experts whose expertise leads to some additional knowledge and those who simply know the field, but do not add extra knowledge to it. There are only $d+2$ possible viewpoints, so, we have $N=d+2$ people who represent N different viewpoints. If m out of these N agree with the statement that "u is a possible value of x", then we take $\mu_X(u)=m/N$ as the degree of belief in this statement. This idea is indeed one of the most widely used ways of estimating the values of a membership function ; see, e.g., Blin et al [9, 8], Dubois et al [16], Section IV.1.d, Klir et al [32].

Average Number of Elements

When the universal set U is finite, we can ask the following natural question: how many elements are possible? For an expert α, the number of possible elements is equal to $|X(\alpha)|$, where by $|S|$, we denote the number of elements in a set S. As a natural measure of size, we can take the arithmetic *average* over all experts. It turns out that this average can be described in terms of the membership function μ_X:

PROPOSITION 6. *Let $\alpha_i=(i+1)/(d+2)$. Then, for every nested set X,*

$$\frac{1}{d+2}\sum_{\alpha\in A}|X(\alpha)|=\sum_{u\in U}\mu_X(u).$$

Historical comment. For fuzzy sets, this number was introduced by Zadeh and called *sigma-count* [81, 82].

The Simplest Case of $d=0$: Tetrvals and Twins

Intervals with intervally uncertain bounds. If experts that were initially unsure make their decision, then the resulting degree of belief in S can take any value from the *interval* $[d(S),pl(S)]$: the value $d(S)$ corresponds to the case when they all decide against S, and the right-hand side $pl(S)$ corresponds to the case when all these experts choose to believe in S. We have already mentioned that the ratio $N(S)/N$ depends on N and is therefore, not a very

good description of the experts' belief. To get a better description, we can use an *interval* $\mathbf{d}(S) = [d^-(S), d^+(S)]$ of possible values (e.g., $d^{\pm}(S) = (N(S) \pm 0.5)/N$. Similarly, we get an interval $\mathbf{pl}(S)$ of possible values of plausibility. As a result, we get a representation of the expert's belief in S by a *pair* of intervals $[\mathbf{d}(S), \mathbf{pl}(S)]$. This representation was proposed in Atanassov et al [2, 4, 3], and called *interval-valued intuitionistic fuzzy logic*. This description is equivalent to the following one: instead of a single interval $\mathbf{d}(S) = [d(S), pl(S)]$ to represent the expert's degree of belief, we now have *two* nested intervals: a more cautious one $\mathbf{D}(S) = [d^-(S), pl^+(S)]$, and a more risky one $\mathbf{d}(S) = [d^+(S), pl^-(S)] \subseteq \mathbf{D}(S)$.

For intervals with intervally uncertain bounds, we have a modified problem:

- For *expert systems*, we have a knowledge base that consists of statements $S_1, ..., S_N$; for a given query Q, we also have an algorithm (obtained by using the inference engine) $f(d_1, ..., d_N)$ that transforms the degrees of belief $d(S_1), ..., d(S_N)$ in the statements from the knowledge base into a degree of belief $d(Q) = f(d(S_1), ..., d(S_N))$ in Q.
- For *intelligent control*, we have an algorithm that transforms these same degrees of belief $d(S_1), ..., d(S_N)$ into the recommended control value $\bar{u} = f(d(S_1), ..., d(S_N))$.

Suppose now that we know only the *nested intervals* $\mathbf{d}(S_1) \subseteq \mathbf{D}(S_1)$, ..., $\mathbf{d}(S_N) \subseteq \mathbf{D}(S_N)$ that describe the expert's degrees of belief. From this information, we want to compute the nested interval $\mathbf{d}(Q) \subseteq \mathbf{D}(Q)$ (or $\mathbf{u} \subseteq \mathbf{U}$) for the result. Here,

$$\mathbf{d}(Q) = f(\mathbf{d}(S_1), ..., \mathbf{d}(S_N)) = \{f(d_1, ..., d_N) | d_i \in \mathbf{d}(S_i)\},$$

$$\mathbf{D}(Q) = f(\mathbf{D}(S_1), ..., \mathbf{D}(S_N)) = \{f(d_1, ..., d_N) | d_i \in \mathbf{D}(S_i)\}$$

(and similarly for $\mathbf{u}$ and $\mathbf{U}$). So, one way to compute the desired nested interval is to apply interval computations techniques twice and compute both intervals. In some case, the function f is monotonic in each of the variables, so, for each α, we can simply compute the values of f at the corresponding endpoints. (For nested intervals in intelligent control, this idea is described in Dubois et al [15].)

Another possibility is to apply algorithms that handle the nested interval as a single object. Such methods were proposed in Gardeñes, Zyuzin, Musaev [19, 89, 90, 45, 46], where the nested interval is called a *twin* or a *tetrval*.

Comment. A similar notion of "uncertainty of systematic uncertainty" has also been proposed in Loo [38] in slightly different terms, but, as shown in Artbauer [1], it is essentially a tetrval. A similar idea of describing uncertainty by two intervals was proposed by Narin'yani [47, 48, 49].

2.5 Nested Sets and Nested Intervals: General Case

In the previous text, we considered the case when the experts could be linearly ordered according to their cautiousness. In real life, an expert can be more cautious in one aspect, but more risky in another, so comparing the experts' riskiness does not necessarily lead to a linear order. We can still define an ordering by saying that an expert β is *riskier* than expert α (and denoting it by $\alpha \le \beta$) if for every quantity x, the set $X(\alpha)$ of possible values of x according to α contains all the values $X(\beta)$ that experts β consider possible (i.e., if $X(\alpha) \supseteq X(\beta)$), but this ordering is no longer linear.

When all experts were linearly ordered, our choice was limited: we could choose one of them and consider whatever this expert says to be true. In the general situation, we have more choices. For example, we can choose two experts α and β whom we consider the most credible, and believe everything that each of these two experts say. Therefore, if for some quantity x, these two experts describe the sets $X(\alpha)$ and $X(\beta)$ of possible values, then according to this description of beliefs, we believe that the actual value of x must belong to the intersection $X(\alpha) \cap X(\beta)$. It is reasonable to describe the resulting belief situation as $\alpha \cap \beta$. On the other hand, we may not completely trust each of these experts individually, but we can believe that one of them is right. For this degree of belief, possible values of x form the set $X(\alpha) \cup X(\beta)$. This situation can be described as $\alpha \cup \beta$.

In general, we can take more complicated descriptions of beliefs: e.g., we can say that we believe something if either expert 1 says it, or experts 2 and 3 both agree on it. The resulting "degrees of beliefs" satisfy the following property: whatever degrees of beliefs α and β we have, we can always come up with the degree of belief $\alpha \cap \beta$ (that corresponds to believing in everything in both α and β), and with the degree of belief $\alpha \cup \beta$ (that corresponds to believing that one of the beliefs α and β was correct). In mathematical terms, the set of possible degrees of belief thus becomes a *lattice*, in which the *union* $\vee$ (the least upper bound) corresponds to $\cap$, and the *meet* $\wedge$ (the greatest lower bound) corresponds to $\cup$.

For the basic quantities, we have $X(\alpha \cap \beta) = X(\alpha) \cap X(\beta)$ and $X(\alpha \cup \beta) = X(\alpha) \cup X(\beta)$. However, these properties are not necessarily true for derivative quantities. For example, assume that the first expert α_1 believes that $x \in X(\alpha_1) = \{-1, 0\}$, and the second expert α_2 believes that $x \in X(\alpha_2) = \{0, 1\}$. Then, if we believe both experts, we can conclude that $x \in X(\alpha_1 \cap \alpha_2) = X(\alpha_1) \cap X(\alpha_2) = \{0\}$. However, for $y = x^2$, this property is no longer true: for the first expert, we have $Y(\alpha_1) = \{0, 1\}$. Similarly, for the second expert, we have $Y(\alpha_2) = \{0, 1\}$. However, if we believe both experts, then $x = 0$, $y = 0$, and hence, $Y(\alpha_1 \cap \alpha_2) = \{0\} \neq Y(\alpha_1) \cup Y(\alpha_2)$.

Similarly to the linearly ordered case, in addition to the degrees of belief that describe combinations of expert opinions, we have an element α_0 for which $X(\alpha_0)$ describes the set of elements that can be described as possible without the necessity to believe any of the experts, and an element α_{d+1} for which $X(\alpha_{d+1})$ describes the elements *proven* to be possible. For every other degree of belief α, $X(\alpha_0) \supseteq X(\alpha) \supseteq X(\alpha_{d+1})$. Therefore, if we define $\alpha_0 \cup \alpha$ and $\alpha_{d+1} \cup \alpha$ as $X(\alpha_0 \cup \alpha) = X(\alpha_0) \cup X(\alpha)$ and $X(\alpha_{d+1} \cup \alpha) = X(\alpha_{d+1}) \cup X(\alpha)$, we will conclude that $\alpha_0 \cup \alpha = \alpha_0$ and $\alpha_{d+1} \cup \alpha = \alpha$. Similarly, $\alpha_0 \cap \alpha = \alpha$ and $\alpha_{d+1} \cap \alpha = \alpha_{d+1}$. In algebraic terms, this means that in the lattice of degrees of belief, elements α_0 and α_{d+1} play the role of 0 and 1. As a result, we arrive at the following definition.

Definition 1′.

- *Let an integer d be fixed. This number will be called the number of experts.*
- *Let a finite lattice A be fixed, with elements α_0 and α_{d+1} that are 0 and 1 (i.e., for which $\alpha_0 \vee \alpha = \alpha_0$, $\alpha_{d+1} \vee \alpha = \alpha$, $\alpha_0 \wedge \alpha = \alpha$, and $\alpha_{d+1} \wedge \alpha = \alpha_{d+1}$). Elements of A will be called degrees of belief.*
- *Let a set U be fixed. This set U will be called the Universal set, or the Universe of discourse.*
- *By a piece of information (or, a nested set), we mean a non-decreasing mapping X from A to the set 2^U of all subsets of U, i.e., a mapping for which if $\alpha \leq \beta$, then $X(\alpha) \supseteq X(\beta)$.*
- *Let u be an element of U, and $\alpha \in A$ be a degree of belief. We say that u belongs to the nested set X with a degree of belief $\geq \alpha$ if $u \in X(\alpha)$.*

We can define intersection and union of nested sets by Definitions 2 and 3. It is easy to check that if X and X' are nested sets, then their union and intersection

are nested sets as well. The result of processing nested sets can be described by Definition 5. Since A is a finite lattice, every set has a supremum, and we can use Definition 6 to define a membership function $\mu_X(u)$. However, in this case, Proposition 5 will no longer be true. For example, in the above example, $Y(\alpha_1) = Y(\alpha_2) = \{0, 1\}$, so, $\mu_Y(1) = \sup\{\alpha_1, \alpha_2\} = \alpha_1 \vee \alpha_2$. So, for $u = 0$ and $\alpha = \alpha_1 \vee \alpha_2 = \alpha_1 \cap \alpha_2$, we have $\mu_Y(u) \geq \alpha$. However, as we have already mentioned, $1 = u \notin Y(\alpha) = Y(\alpha_1 \vee \alpha_2) = \{0\}$.

Historical comment. Membership functions with values in a lattice were originally proposed in Goguen [22, 23].

In such cases, an ideal way to store the experts' opinion is to store the conclusions of all the experts. However, in many cases, this takes too much storage space. To reduce the storage space, we may want to follow the idea applied above: forget about who said what, and only store the number of experts who have voted for a given statement. This approach was used by Kohout et al (see, e.g., [34] and references therein).

The problem with this approach is as follows: When we store the complete information about the experts' opinion, then for every two statements A and B, from the information about A and the information about B, we can determine the experts' opinion about B. However, if we only know the number of experts who believe in A, and the number of experts who believe in B, then we cannot determine uniquely the number of experts who believe in $A\&B$. For example, if half of the the experts believe in A, and half believe in B, then it could be that the same experts who believe in A also believe in B; in this case, exactly half of the experts believe in $A\&B$. It can also happen that none of the experts who believe in A believe in B, in which case, none of the experts believe in A and B at the same time. So, instead of a single value of $d(A\&B)$, we get an *interval* of possible values of $d(A\&B)$. This situation is described in the surveys Nguyen et al [54, 55] (see also references therein).

3 OTHER PROBLEMS WHERE NESTED SETS AND NESTED INTERVALS CAN BE USED: IDENTIFICATION, OPTIMIZATION, CONTROL, AND DECISION MAKING

In the above text, we mainly concentrated on the problem of estimating accuracy of the results of indirect measurements. In this section, we will show how nested sets and nested intervals can be (and are) used in other application problems: identification, optimization, control, and decision making.

3.1 Identification Problem for Nested Sets and Nested Intervals

In general, the identification problem can be formulated as follows: we know that a physical quantity y depends on the quantities $x_1, ..., x_n$; we want to find the dependency $y = f(x_1, ..., x_n)$. In order to find the dependency, we measure x_i and y in several situations. As a result, we get finitely many tuples $(x_1^{(k)}, ..., x_n^{(k)}, y^{(k)})$, $1 \le k \le N$ (they are usually called *patterns*), and we must find a function f that satisfies N conditions $y^{(k)} = f(x_1^{(k)}, ..., x_n^{(k)})$. There are infinitely many functions that satisfy these conditions. In order to choose one of them, we must have some additional knowledge about the function f (i.e., we must restrict f to some finite-dimensional class). Often, this class consists of linear functions or, in a slightly more general setting, functions that are linear combinations of known ones, i.e., functions of the type $f(x_1, ..., x_n) = c_1\phi_1(\vec{x}) + ... + c_m\phi_m(\vec{x})$, where the base functions ϕ_j are given, and the coefficients c_j have to be determined from the experiments.

If for every measurement, we only know the interval of possible values of error, then we get an interval identification problem. If, in addition to the guaranteed estimates, we have experts' estimates of these same errors, then the possible values of x_i and y in each measurement are characterized by *nested sets* (in particular, by nested intervals). In this case, we have a nested set (interval) identification problem (see, e.g., Tamura et al [62]): we have nested intervals $X_i^{(k)}$ and $Y^{(k)}$, and we want to find a function f for which $f(X_1^{(k)}, ..., X_n^{(k)}) = Y^{(k)}$ for some mapping f from tuples of nested intervals into nested intervals. A nested interval X can be described by the bounds $x \pm (\alpha)$ of its intervals $X(\alpha)$,

$\alpha \in A$. So, in general, each value $y^{\pm}(\alpha)$ is a function of $2n(d+2)$ variables $x_i^{\pm}(\beta)$ for all $\beta \in A$.

This complicated regression problem can be simplified, if we take into consideration that the nested interval identification problem makes sense only if we use a single measuring instrument for each of the variables x_i and y: for example, if we use two different measuring instruments (with different accuracies) to measure y, then, even if we know the results of measuring x_i, we cannot predict the interval $Y(\alpha)$, because we do not know which of the instruments was used.

In case a single instrument is used, for every α and β, the larger $x^-(\alpha)$, the larger the measurement result $\tilde{x}$ and therefore, the larger $x^{\pm}(\beta)$. So, for every quantity x, there is a 1-1 correspondence between the variables $x^-(\alpha)$ and $x^{\pm}(\beta)$. In view of this correspondence, there is no need to consider all $2(d+2)$ values $x_1^{\pm}(\alpha_i)$ as inputs for the function f: one of these variables will be sufficient.

It is convenient, when predicting $y^{\pm}(\alpha)$, to choose the variables $x_i^{\pm}(\alpha)$ that correspond to this same degree of belief α and to this same bound (left if $-$, right if $+$). As a result, we have the numerical identification problem $y^{\pm}(\alpha) = f^{\pm}(\alpha, x_1^{\pm}(\alpha), ..., x_n^{\pm}(\alpha))$. In particular, if we know that the dependency between y and x_i is linear, it is natural to choose each function $f^{\pm}(\alpha)$ to be also linear in its variables, i.e., $y^{\pm}(\alpha) = c_0^{\pm}(\alpha) + c_1^{\pm}(\alpha) \cdot x_1^{\pm}(\alpha) + ... + c_n^{\pm}(\alpha) \cdot x_n^{\pm}(\alpha)$ for some coefficients $c_i^{\pm}(\alpha)$. So, we have $2(d+2)$ numerical regression problems: two (corresponding to $+$ and $-$) for each $\alpha \in A$. Similarly to the main problem, we have two possibilities here: One possibility is to solve the resulting regression problems one by one. Another is to try to combine them together. The values $c_i^{\pm}(\alpha)$ resemble the bounds $x_i^{\pm}(\alpha)$ that form a nested interval. However, for the coefficients, it is not necessarily true that $c_i^-(\alpha) \le c_i^+(\alpha)$ or that the resulting intervals $[c_i^-(\alpha), c_i^+(\alpha)]$ are decreasing in α. So, to describe the coefficients $c_i^{\pm}(\alpha)$ as a single object, we must describe some generalization of the notion of the nested set. Such a generalization was proposed in [62].

3.2 Optimization for Nested Sets and Nested Intervals

Suppose that we must optimize the function $f(\vec{x}, \vec{c})$ under the conditions that $g_i(\vec{x}, \vec{c}) \le 0$. Here, $\vec{x} = (x_1, ..., x_n)$ are the desired variables, and $\vec{c} = (c_1, ..., c_m)$ are parameters that are known with uncertainty (i.e., only nested sets C_i are known). Then, for each degree of belief α, we believe that $c_i \in C_i(\alpha)$.

If we fix $\vec{c}$, we get a constrained optimization problem; let us define the set of its solutions $(x_1, ..., x_n)$ by $X(\vec{c})$. (This set can consists of a single solution, or it can contain several, if the optimum is not attained at a single point $\vec{x}$.)

For each degree of belief α, we can define the set $S(\alpha)$ of possible solutions to the optimization problem as the set of all solutions $\vec{x}$ that are possible under this degree of belief; in mathematical terms, $S(\alpha)$ is the union of the solution sets $X(\vec{c})$ for all $\vec{c}$ for which $c_i \in C_i(\alpha)$. These sets $S(\alpha)$ form a *nested set* that can thus be viewed as a nested solution of the optimization problem.

Optimization problems for fuzzy (=nested) sets have been first formulated in Bellman et al [5].

Optimization has been done using intervals for a long time (see, e.g., a survey by Kearfott [30]). The majority of the methods of nested set optimization are, in effect, reasonably straightforward generalizations of these interval optimization techniques.

For a detailed exposition of the modern state of research in this field, the see, e.g., Zimmermann, Orlovsky et al [83, 56, 84, 57, 85, 86, 87, 88].

3.3 Control and Decision Making

Decision Making with Finitely Many Alternatives

Suppose that we must choose between two or finitely many alternatives based on a single criterion (i.e., cost or reliability). To make a choice, we must compute the value of this criterion for each parameter, and compare the resulting values. The desired value of the criterion depends on some parameters that we must measure. Due to uncertainty of measurements, for each alternative, we get not a single value of the criterion, but an *interval* of possible values. If for alternative a the resulting interval $[a^-, a^+]$ is guaranteed to be larger than the interval $[b^-, b^+]$ that corresponds to the parameter b (i.e., if $a^- \geq b^+$), then we definitely prefer a to b. But what should be done if these two intervals intersect, and therefore, a can be better or worse than b? In some cases, in addition to the guaranteed interval, we have expert estimates that lead to narrower intervals. In these cases, Loui [39] proposes the following: if guaranteed estimates do not enable us to make a unique decision, let us try to make the decision based on the intervals provided by the first (more cautious) expert α_1; if this still does not work, let us use the interval provided by the next expert α_2, etc, until

finally, we find an expert whose interval is narrow enough for the decision to be made.

Decision Making: General Case

In control and decision making, we may look for the best decision. For example, we may look for a decision u for which $J(x,u) \to \max$ for some objective function J, where x is a parameter (or parameters) that we measure. Since measurements are not absolutely precise, we will end up with the interval $X(\alpha_0)$ of possible values of x. For each $x \in X(\alpha_0)$, we can find the value u that is optimal for this particular x. The values corresponding to different $x \in X(\alpha_0)$ form a *set* of possible controls $U(\alpha_0)$. In many cases, this set is an interval, but sometimes, its structure is more complicated.

An expert $\alpha \in A$ can believe that a smaller set $X(\alpha)$ of values of x is possible. This expert will thus come up with a set $U(\alpha)$ that is a subset of $U(\alpha_0)$. So, we have a *nested set* of possible values of control. We need to choose a single value of u. For decision making, we can simply present the nested set to the decision maker and let the decision maker choose. For automatic control, we must automatically select a number from the set. In fuzzy set language, the transition from a nested (fuzzy set) to a number is called *defuzzification*.

There are two basic approaches to defuzzification:

- First, since we need to select a single value, it makes sense to go with the bravest of the experts (the one whose choice is the narrowest); if this expert has already selected a single control value, we will choose it, otherwise, we will choose from the set $X(\alpha)$ provided by this expert.
- The second approach is to take into the consideration the opinion of all the experts.

Let us analyze these two approaches. We will start with the first approach. Let us assume (for simplicity) that the set from which we must choose is an interval. So, the question becomes: how to choose from an interval? We must describe a *choice function* s that transforms an interval $\mathbf{a} = [a^-, a^+]$ into a number $a = s(\mathbf{a}) \in \mathbf{a}$. This function must satisfy the following natural properties. First, if we change the starting point for measuring a (time and temperature are good examples where such a change is possible), then all the measured values are shifted ($x \to x + c$ for a fixed c). This is a formal change

that does not affect our knowledge, like a change from Kelvin to centigrade in measuring temperature. Therefore, it is natural to assume that the choice provided by the function s should not change after a change; in other words, that if we first make a choice (i.e., go from $\mathbf{a}$ to $s(\mathbf{a})$) and then change the starting point (i.e., go from $s(\mathbf{a})$ to $s(\mathbf{a}) + c$), the result must be the same as if we first change the scale (from $[a^-, a^+]$ to $[a^- + c, a^+ + c]$) and then make a choice: i.e., $s([a^- + c, a^+ + c]) = s([a^-, a^+]) + c$.

Similarly, If we change the unit in which we measure a, the choice must not change. If we change the unit (e.g, go from inches to centimeters), then, the numerical values of this quantity change as $x \to \lambda x$ for some $\lambda > 0$ (= the ratio of the old and the new units). So, we arrive at the formula $s([\lambda \cdot a^-, \lambda \cdot a^+]) = \lambda \cdot s([a^-, a^+])$.

It is also natural to assume that the choice function should be invariant w.r.t. changing the sign of a: e.g., if a is the angle on which to steer the wheel, then the sign is just a convention, and the result should not depend on whether we decided to call clockwise angles positive or negative. So, we arrive at the following additional invariance: $s([-a^+, -a^-]) = -s([a^-, a^+])$. Combining these three conditions together, we obtain the following result:

Definition 7. *By a choice function, we mean a function s that maps every interval $[a^-, a^+]$ into a point from that interval, and that has the following properties for every interval and for every c and $\lambda > 0$:*

- $s([a^- + c, a^+ + c]) = s([a^-, a^+]) + c$ *(shift-invariance);*
- $s([\lambda \cdot a^-, \lambda \cdot a^+]) = \lambda \cdot s([a^-, a^+])$ *(unit-invariance);*
- $s([-a^+, -a^-]) = -s([a^-, a^+])$ *(symmetry).*

PROPOSITION 7. *Every choice function has the form*

$$s([a^-, a^+]) = \frac{a^- + a^+}{2}.$$

So, we take the *midpoint* of the interval of possibly best decisions. In fuzzy terms, this defuzzification method (midpoint of the smallest interval that corresponds to the largest α) is called *mean of maxima* . This approach can be interpreted as follows: we consider all possibly best decisions, and we take the

average of them (with equal weight). In mathematical terms, in the discrete case, this average would be equal to the ratio

$$\left(\sum_{x\in X} x\right) / \left(\sum_{x\in X} 1\right),$$

and in the continuous case, it is equal to the ratio

$$\left(\int_X x dx\right) / \left(\int_X dx\right).$$

If we have *several* experts who make different predictions, then in this averaging process, we can ask every expert to write down all the values that this expert considers possible, and average them. In this new averaging, every value is counted as many times as many experts consider it possible. In other words, we propose the following ratio:

$$\left(\sum_{\alpha\in A}\int_{X(\alpha)} x dx\right) / \left(\sum_{\alpha\in A}\int_{X(\alpha)} dx\right).$$

This ratio can be expressed in terms of the membership function $\mu_X(x)$:

PROPOSITION 8.

$$\left(\sum_{\alpha\in A}\int_{X(\alpha)} x dx\right) / \left(\sum_{\alpha\in A}\int_{X(\alpha)} dx\right) = \left(\int x\cdot\mu_X(x)dx\right) / \left(\int \mu_X(x)dx\right).$$

This formula is called *center of gravity defuzzification* .

These two formulas work well for *nested intervals*, but they do not always help for nested sets that are not intervals. For example, if the car is approaching an obstacle, then we can either turn to the left, or turn to the right. If the situation is absolutely symmetric, then the possible set of controls is, e.g., a union of two intervals $[-\Delta,-\delta]\cup[\delta,\Delta]$. Because of symmetry, the average of this set is 0, so, our choice will lead us straight ahead to the obstacle.

When the set from which we choose is a union of several intervals, Yen et al [75, 76, 77] propose taking the largest interval, and then taking the midpoint of this largest interval. For nested sets for which $X(\alpha)$ is a union of intervals, we can choose an interval that captures the most votes (i.e., for which $\int \mu_X(x)dx$ is the largest), and apply the formula described above (center-of-gravity defuzzification) to that interval only.

4 APPLICATIONS OF NESTED SETS AND NESTED INTERVALS

In this section, we survey the existing applications of nested sets and nested intervals.

4.1 Computing the Nested Interval for the Result of Data Processing

Mathematical comment. In this subsection, we survey applications of nested sets and nested intervals to the problem of obtaining the *true range* of a function, the problem that we described in the Introduction as one of the main problems of data processing. In mathematical terms, for each degree of belief α, we have a function $f(x_1, ..., x_n)$ and n intervals $X_i(\alpha)$, and we want to find the interval $Y(\alpha) = [y^-(\alpha), y^+(\alpha)]$ of the possible values of $f(x_1, ..., x_n)$ for $x_i \in X_i(\alpha)$. From the mathematical viewpoint, this problem is effectively an *optimization problem*: indeed:

- $y^-(\alpha)$ is the solution of the constrained optimization problem: $f(x_1, ..., x_n) \to \min$ under the constraints $x_i \in X_i(\alpha)$.
- $y^+(\alpha)$ is the solution of a similar constrained optimization problem: $f(x_1, ..., x_n) \to \max$ under the constraints $x_i \in X_i(\alpha)$.

Therefore, these applications can be viewed as particular cases of *applied optimization*.

Case study: seismic engineering (Dong et al [13]). As we have already mentioned in the introduction, earthquakes are extremely difficult to predict. Because of that, if we only use the equations and the measured data x_i, we get very wide intervals of possible magnitudes and frequencies y, intervals that are known to be much wider than the interval of actual values. So, if we use these wide intervals $[y^-, y^+]$ to design a building that is guaranteed to withstand a typical earthquake, these buildings will be unnecessarily (and often unrealistically) expensive. To make the requirements more realistic, in addition to guaranteed measurement results, expert estimates are normally used. For every parameter x_i (e.g., for the frequency of earthquakes), estimates by different experts form a nested interval (sometimes, a nested set). In [13], interval methods are used for an expert-by-expert processing of these nested intervals.

Namely, to estimate the range of f on a box $X_1(\alpha) \times ... \times X_n(\alpha)$, the box is subdivided into several sub-boxes small enough that each sub-box contains at most one local extremum point; for each sub-box, the range is estimated. Then, the union of the estimates for sub-box ranges is taken as an estimate for the total range. On each sub-box, the maximum of f is attained either at an interior point (in which case this point is a local extremum of f), or on one of the sides. For each side, we can repeat a similar procedure, and end up with a conclusion that the maximum is attained either at one of the corners, or at a local extremum point of one of the sides of smaller dimension. There are at most $\approx 2^n$ vertices and local extremum points in each sub-box, so, for small n, we can simply enumerate them all. The resulting nested intervals Y are presented to the decision makers.

Case study: AIDS epidemics (Berleant et al [6]). This is another example of usefulness of nested sets presented in the Introduction. Equations that describe how AIDS is transferred are known. To solve these equations, we must know the initial conditions and the the numerical values of the parameters (like the death rate among AIDS patients, etc). The majority of these parameters can be determined more or less accurately from known statistical data. However, there is one important parameter about which it is very difficult to get a reliable statistical estimate: the number of intravenous drug abusers. So, in the absence of a reliable statistical estimate, we may use estimates of experts. Different experts can give different intervals of possible values, so we have nested intervals for describing the expert knowledge (that can be described by a membership function). Applying interval computations to intervals corresponding to several levels, we get a nested interval for the predicted number of AIDS patients in the year 2000 [6]. Due to the serious input uncertainty, all level intervals are so wide that we cannot simply neglect quadratic terms and apply linearization techniques; therefore, we have to use interval computation techniques.

To get a better estimate, it is reasonable to *combine* these expert estimates with the statistical estimates. A natural way to combine expert knowledge with statistical knowledge is to represent the statistical information (that is given in terms of a probability density function) in terms of nested intervals; then, we will simply take an intersection of the two nested intervals that represent expert and statistical knowledge. This combination can be done as follows: crudely speaking, $X(\alpha)$ describes the interval in which $\alpha-$th part of the experts believe. So, to get a similar interval from the probability density function ρ, we can take the set $S = \{u|\rho(u) \geq p_0\}$ for some p_0 for which the probability to belong to S is exactly α. This approach has lead to a reasonable prediction [6].

Case study: engineering design (Yang et al [69]). In design, we would like to predict the properties of a designed system without actually designing it in hardware. So, we need to estimate the values $y = f(x_1, ..., x_n)$ for different functions f and for different nested intervals x_i. In manufacturing measurements, the accuracy is limited by cost considerations, so the intervals $X_i(\alpha)$ are not very small, and therefore, we cannot neglect quadratic terms and use linearization techniques. Instead, methods of interval computations are used.

Case study: decision analysis and risk analysis (Dong et al [14]). Usually, we have several criteria $c_1, ..., c_m$ to choose between several alternatives $a_1, ..., a_n$ (e.g., when we are choosing a car, we may opt for speed, or for fuel efficiency, or for cost, etc). We may choose the alternative a_i for which the first criterion c_1 takes the largest possible value, i.e., for which $c_1(a_i) \to \max_i$, or we can take the alternative for which $c_2(a_i) \to \max_i$, etc. A natural way to handle this multi-criteria situation is to assign to every alternative a *weighted average* of the values $c_1(a_i), ..., c_i(a_i)$, i.e., the value $r_i = (w_1 c_1(a_i) + ... + w_n c_1(a_n))/(w_1 + ... + w_n)$. The values $c_j(a_i)$ are obtained from measurements, so they are represented by nested intervals. The weights are also obtained from analysis of experiments (or simply by asking experts), so the weights are also nested intervals. So here, the x_i are values $x_1 = c_1(a_i), ..., x_n = c_1(a_n), x_{n+1} = w_1, ..., x_{2n} = w_n$, and $f(x_1, ..., x_{2n}) = (x_1 x_{n+1} + x_2 x_{n+2} + ... + x_n x_{2n})/(x_{n+1} + ... + x_{2n})$. (Sometimes, a more complicated combination function f is used.) The nested intervals for c_{ji} and w_i are reasonably large (especially in the design stage), so we cannot apply linearization techniques. In [14], a degree-by-degree approach is used. Under many reasonable assumptions, the function f is monotonic. For a monotonic function f, for every α, the maximum and minimum for $x_i \in X(\alpha) = [x^-(\alpha), x^+(\alpha)]$ are attained for $x_i = x^{\pm}(\alpha)$. Therefore, for every α, it is sufficient to consider the values of f for 2^{2n} "corners" $(x_1^{\pm}, ...x_{2n}^{\pm})$. For small n, this a reasonable procedure, much faster than other known numerical techniques [14].

If we know that a function f is monotonically increasing in each variable, then we can only consider two points instead of 2^n: $f(x_1^-(\alpha), ..., x_n^-(\alpha))$ is the lower bound of the desired interval $Y(\alpha)$, and $f(x_1^+(\alpha), ..., x_n^+(\alpha))$ is the upper bound for this interval (Dubois et al [16]). Another case when a simpler algorithm is possible is when f is a convex function (see Schmucker [59]).

Comment. Several methods and examples of processing fuzzy inputs are discussed in Mareŝ [43].

4.2 Optimization

Mathematical comment. We have already mentioned that, in the above examples, although we start with a practical problem that is not formulated in terms of optimization, the problem is equivalent to an optimization problem. In other cases, the applied problem itself is already formulated as an optimization problem. Examples of such applications are given in this subsection.

Case study: agriculture (Shaocheng [60]). When we plan to run a chicken farm, we must know how much money will be required to feed the chickens. Experts can tell us how many calories, proteins, vitamins, etc, a chicken needs per day; how many of these are contained in each type of available chicken food (e.g., soya, millet, etc), and what is the cost per unit for each type of food. We must find the cost of the cheapest combination that provides for the chickens. In mathematical terms, if we denote the total number of available types of food by n, and the total number of necessary food ingredients that the menu must contain by m, then the problem is: find the amounts x_i, $1 \le i \le n$, of different components that minimize the total cost $c = c_1x_1 + ... + c_nx_n$ under the conditions that the total amount of every ingredient is sufficient, i.e., that the total amount $a_{j1}x_1 + ... + a_{jn}x_n$ of this ingredient in the resulting mixture must exceed the chicken's demand d_j: $a_{j1}x_1 + ... + a_{jn}x_n \ge d_j$ for all j. At any given moment of time, we will know the exact current values of the parameters c_i, a_{ij}, and d_j, and we will be able to find the optimal values of x_i and the corresponding cost c by solving a linear programming problem. However, in the planning stage, we can only have *interval* predictions of the costs and of, say, protein contents of the future crop. In other words, from experts, we can get *intervals* $\mathbf{c}_i = [c_i^-, c_i^+]$, $\mathbf{a}_{ij} = [a_{ij}^-, a_{ij}^+]$, $\mathbf{d}_j = [d_j^-, d_j^+]$ of possible values of the corresponding parameters. In this case, we can only get an *interval* estimate for the cost. From the mathematical viewpoint, we must solve the *interval linear programming* problem: describe the set of all possible costs for all linear programming problems that can be obtained by selecting numbers from the given intervals. For this particular problem, the desired interval $\mathbf{c} = [c^-, c^+]$ is easy to describe: the lowest overall cost corresponds to when the cost of the components is lowest, the contents is highest, and the demand is the lowest possible. In other words, c^- is the solution of the linear programming problem $c^- = c_1^- + ... + c_n^-x_n \to \min$ subject to the constraint $a_{j1}^+x_1 + ... + a_{jn}^+x_n \ge d_j^-$. Similarly, the largest possible optimal cost is a solution of the linear programming problem $c^+ = c_1^+ + ... + c_n^+x_n \to \min$ subject to the constraint $a_{j1}^-x_1 + ... + a_{jn}^-x_n \ge d_j^+$. The problem with this approach is that some experts may give more cautious interval estimates while some give riskier (narrower) interval estimates. In other words, the knowledge of these experts

can be described by *nested* intervals. In this case, the solution is also a nested interval, with levels that correspond to different experts. In [60], both expert-by-expert and membership function approaches are used to describe the nested interval that corresponds to the solution.

Similar mathematical problems (leading to nested interval linear programming) arise in several different applications that are described in the remaining part of this subsection. (We highlighted the chicken feed application, because in the corresponding paper, the relationship between the applied methods and interval computations are explicitly explained; in the other papers, this relationship may be obscured. For a general mathematical formalism, see, e.g., Yazenin [70, 71, 72, 73, 74] and Lodwick [37].)

Case study: air pollution (Sommer et al [61]): Nested interval linear programming is used to determine how to regulate different manufacturing processes with different pollutants so that the cost is minimized under the condition that the overall concentration of different pollutants does not exceed the given limits.

Case study: media selection (Wiedy et al [67]): Each medium (newspaper ads, TV, radio, etc) targets a special audience. Nested interval linear programming describes how to target the largest number of people at a minimal cost.

Case study: scheduling (Deporter et al [12]): The parameters to select are starting times of different stages of the large project (and/or the percentage of resources that are allocated to different stages of the project in different moments in time). The restrictions describe the necessity to finish some stages so we can start others, and limited resources that prevent us from doing too many stages at the same time. The goal is to minimize the total time of the project. Uncertainty is caused by the impossibility to predict exactly the duration of each stage, and the future amount of resources.

Case study: reliability problems (Sasaki et al [58]): This involves choosing the cheapest design that guarantees the desired reliability of different regimes. Here, the parameters x_i describe the reliability of different components (namely, the average number of failures per time unit). The overall reliability of different regimes can be described as a linear function of component reliabilities. The cost of the corresponding components can also be described as a linear function of x_i (if we can neglect the terms that are quadratic in x_i).

Case study: design of a manufacturing plant (Tsai [65]). Here, the parameters x_i describe the number of manufacturing units of different types. We need sufficiently many to be able to successfully perform several different types of operations (i.e., the total ability $a_{j1}x_1 + ... + a_{jn}x_n$ to perform $j-$th operation must exceed the desired number d_j), and the goal is to minimize the total cost.

Case study: controlling a highway network (Bit et al [7]). Here, the control parameters $\vec{x} = (x_1, ..., x_n)$ are car flows sent to different routes. The goal is to increase the total car flow. Restrictions describe the finite capacity of the streets and roads. The interval and nested interval uncertainty is caused by such difficult-to-predict parameters such as sudden weather changes, unexpected soccer results, or an amazing TV chase that glues everyone to their homes.

4.3 Control and Decision Making

Case study: intelligent control. Application of expert knowledge to control problems is called *intelligent control* , or *fuzzy control* (The latter term is due to the fact that fuzzy representation is most often used in control situations).

The expert system approach leads us a nested set for describing possible values of control. To use this nested set, we must choose a single point. The methods of choosing a point that we have described above (mean-of-maximum and center-of-gravity) are indeed the ones that are most frequently used in intelligent control. These approaches were initiated in Mamdani [40, 41], and have since been applied to cement kilns, appliances, subways, helicopters, robots, etc. For a latest list of applications of intelligent control , see, e.g., Kandel et al [28] and Kruse [35].

Case study: robotics (Bojadziev et al [10, 11]). Some robots are designed for the manufacturing floor, where all tasks and conditions are well-defined. But other robots (e.g., robots for space exploration or for handling emergency situations in nuclear power stations) are designed for environments whose properties are not completely known. In this case in addition to the guaranteed intervals that contain possible values, we have narrower intervals that correspond to expert estimates for the unknown parameters. Thus, we have a nested interval. To compare different control strategies, we can compare their quality on intervals $X(\alpha)$ that correspond to different degrees of belief. For example, if we are interested in the worst-case behavior, i.e., in the smallest

possible value of the objective function $J(x)$, we can report to the decision maker the values

$$J(\alpha) = \min_{x \in X(\alpha)} J(x)$$

that correspond to different $\alpha \in A$.

4.4 Expert Systems

In Uehara et al [66], nested intervals describe experts' knowledge about different values.

In Zadeh, Nazaraki et al [81, 82, 50], nested sets are used to interpret expert informal statements of the type $Q(X \to Y)$ for some informal "quantifier" Q and properties X and Y. For example, if $Q =$"many", $X =$"small", and $Y =$"strong", we get a statement "many small people are strong". We would like to interpret this statement as meaning that among people who have the property X, many also have a property Y, i.e., that the ratio $|X \cap Y|/|X|$, where $|X|$ denotes the number of people that satisfy the property X, is large enough to correspond to "many". To formalize this idea, the authors of the cited papers interpret $|X|$ and $|X \cap Y|$ as sigma-counts and choose an interval $\mathbf{q} = [q^-, q^+]$ of values of the ratio that corresponds to "many" (or to any other "quantifier" Q). Then, $Q(X \to Y)$ is interpreted as $|X \cap Y|/|X| \in \mathbf{q}$. This idea leads to reasonable description of our common-sense reasoning.

APPENDIX A

PROOFS

Proof of Proposition 1. Since the function $\alpha \to \alpha^\perp$ is strictly decreasing, it is 1-1, i.e., different α lead to different $\alpha^\perp$. Hence, this function has exactly $d+2$ different values for $d+2$ inputs $\alpha_j \in A$. But the set A has exactly $d+2$ elements. Therefore, each element from A is an image of this map. Due to monotonicity, the values $\alpha_0^\perp, ..., \alpha_{d+1}^\perp$ must go from α_{d+1} to α_0. If we skip one of the values from the sequence $\alpha_{d+1}, \alpha_d, ..., \alpha_0$, then we have fewer than

$d+2$ elements in the image. So, when α passes through the values $\alpha_0, ..., \alpha_{d+1}$, $\alpha^\perp$ must pass through the values $\alpha_{d+1}, ..., \alpha_0$. In other words, $\alpha_j^\perp = \alpha_{d+1-j}$. Q.E.D.

Proof of Propositions 2 and 5 is straightforward.

Proof of Proposition 3. Let us first prove that $\mu_{X\cap X'}(u) \geq \min(\mu_X(u), \mu_{X'}(u))$. Indeed, let us denote $\mu_X(u)$ by α, $\mu_{X'}(u)$ by α', and $\mu_{X\cap X'}(u)$ by $\alpha_\cap$. Without loss of generality, we can assume that $\alpha \leq \alpha'$. By definition of $\mu_X(u)$ as the largest value α from A for which $u \in X(\alpha)$, we conclude that $u \in X(\alpha)$ and, similarly, that $u \in X'(\alpha')$. Since $\alpha \leq \alpha'$, and the sequence of sets $X(\alpha)$ is decreasing in α, we can conclude that $X'(\alpha') \subseteq X'(\alpha)$. Since $u \in X'(\alpha')$, we can conclude that $u \in X'(\alpha)$.

So, $u \in X(\alpha)$ and $u \in X'(\alpha)$, hence, $u \in X(\alpha) \cap X'(\alpha)$. By definition of an intersection of nested sets, this means that $u \in (X \cap X')(\alpha)$. By definition of μ, this means that $\alpha_\cap = \sup\{\alpha | u \in X(\alpha)\} \geq \alpha = \min(\alpha, \alpha')$. So, $\mu_{X\cap X'}(u) \geq \min(\mu_X(u), \mu_{X'}(u))$.

Let us now prove the inverse inequality $\mu_{X\cap X'}(u) \leq \min(\mu_X(u), \mu_{X'}(u))$. Indeed, by definition of μ, we have $u \in (X \cap X')(\alpha_\cap) = X(\alpha_\cap) \cap X'(\alpha_\cap)$. Therefore, $u \in X(\alpha_\cap)$ and $u \in X'(\alpha_\cap)$. From $u \in X(\alpha_\cap)$ and the definition of μ, we conclude that $\alpha_\cap \leq \alpha = \sup\{\alpha | u \in X(\alpha)\}$. Similarly, $\alpha_\cap \leq \alpha'$. Therefore, $\alpha_\cap \leq \min(\alpha, \alpha')$, i.e., $\mu_{X\cap X'}(u) \leq \min(\mu_X(u), \mu_{X'}(u))$.

From $\mu_{X\cap X'}(u) \leq \min(\mu_X(u), \mu_{X'}(u))$ and $\mu_{X\cap X'}(u) \geq \min(\mu_X(u), \mu_{X'}(u))$, we conclude that $\mu_{X\cap X'}(u) = \min(\mu_X(u), \mu_{X'}(u))$.

The proof for $\cup$ is similar. Indeed, let us denote $\alpha = \mu_X(u)$, $\alpha' = \mu_{X'}(u)$, and $\alpha_\cup = \mu_{X\cup X'}(u)$, and let us assume (without losing generality) that $\alpha \leq \alpha'$. Then, from the definition of α', it follows that $u \in X'(\alpha')$, hence, $u \in X(\alpha') \cup X'(\alpha') = (X \cup X')(\alpha')$, and $\alpha_\cup = \sup\{u | u \in (X \cup X')(\alpha)\} \geq \alpha' = \max(\alpha, \alpha')$.

The inverse inequality can be proven in a similar manner: from the definition of $\alpha_\cup$, it follows that $u \in (X \cup X')(\alpha_\cup) = X(\alpha_\cup) \cup X'(\alpha_\cup)$. Since u belongs to the union of two sets, it must belong to one of them. If $u \in X(\alpha_\cup)$, then $\alpha_\cup \leq \alpha$. If $u \in X'(\alpha_\cup)$, then $\alpha_\cup \leq \alpha'$. In both cases, $\alpha_\cup \leq \max(\alpha, \alpha')$.

So, $\alpha_\cup \leq \max(\alpha, \alpha')$ and $\alpha_\cup \geq \max(\alpha, \alpha')$, hence, $\alpha_\cup \leq \max(\alpha, \alpha')$, i.e., $\mu_{X\cup X'}(u) = \max(\mu_X(u), \mu_{X'}(u))$.

Finally, let us prove that $\mu_{X^\perp}(u) = f_\neg(\mu_X(u))$, where $f_\neg(\alpha_j) = \alpha_{d-j}$. Indeed, by definition of $X^\perp$, $u \in X^\perp(\alpha) \leftrightarrow u \notin X(\alpha^\perp)$, where $\alpha_i^\perp = \alpha_{d+1-i}$. So, $u \in X^\perp(\alpha_i) \leftrightarrow u \notin X(\alpha_{d+1-i})$. Hence, $\mu_{X^\perp}(u) = \sup\{\alpha_i | u \in X^\perp(\alpha_i)\} = \sup\{\alpha_i | u \notin X(\alpha_{d+1-i})\}$. Since α_i is a increasing sequence, we can write that $\mu_{X^\perp}(\alpha_i) = \alpha_m$, where $m = \max\{i | u \notin X(\alpha_{d+1-i})\}$. This formula can be simplified if we introduce a new variable $j = d - i$. Then, the largest i corresponds to the smallest value of j, i.e.,

$$m = \max\{i | u \notin X(\alpha_{d+1-i})\} = d - \min\{j | u \notin X(\alpha_{j+1})\}.$$

By definition, $\mu_X(u)$ is α_j for the largest j for which $u \in X(\alpha_j)$. For $\alpha = \alpha_{j+1}, \alpha_{j+2}, ...$, we have $u \notin X(\alpha)$. Therefore, the smallest α for which $u \notin X(\alpha)$, is α_{j+1}, where $\alpha_j = \mu_X(u)$. Hence, $\min\{j | u \notin X(\alpha_{j+1})\}$ is exactly the j for which $\alpha_j = \mu_X(u)$. Hence, $\mu_{X^\perp}(u) = \alpha_{d-j} = f_\neg(\alpha_j) = f_\neg(\mu_X(u))$. Q.E.D.

Proof of Proposition 4. Let us first prove that the left-hand side (LHS) of the desired equality is smaller than or equal to its right-hand side (RHS). Indeed, from $\alpha = \mu_{f(X_1,...,X_n)}(u)$, it follows that $u \in f(X_1, ..., X_n)(\alpha)$. By Definition 5, this means that $u \in f(X_1(\alpha), ..., X_n(\alpha))$, i.e., that $u = f(u_1, ..., u_n)$ for some $u_i \in X_i(\alpha)$. From $u_i \in X_i(\alpha)$, it follows that $\mu_{X_i}(u_i) \geq \alpha$. Therefore, $\min(\mu_{X_1}(u_1), ..., \mu_{X_n}(u_n)) \geq \alpha$, and hence,

$$\max_{u_1,...,u_n : f(u_1,...,u_n)=u} \min(\mu_{X_1}(u_1), ..., \mu_{X_n}(u_n)) \geq \alpha = \mu_{f(X_1,...,X_n)}(u),$$

i.e., RHS≥LHS.

For the converse, let α be equal to the RHS. Since A is a finite set, the supremum is attained for some u_i, so, there exist $u_1, ..., u_n$ for which $u = f(u_1, ..., u_n)$ and $\min(\mu_{X_1}(u_1), ..., \mu_{X_n}(u_n)) = \alpha$. Hence, $\alpha_i = \mu_{X_i}(u_i) \geq \alpha$ for all i. By definition of μ, we have $u_i \in X_i(\alpha_i)$. Since $X(\alpha)$ is a decreasing sequence of sets, we have $u_i \in X_i(\alpha)$. Hence, $u = f(u_1, ..., u_n) \in f(X_1(\alpha), ..., X_n(\alpha)) = (f(X_1, ..., X_n))(\alpha)$. Therefore, this α is smaller than or equal to the largest of all α for which $u \in (f(X_1, ..., X_n))(\alpha)$, i.e., $\alpha \leq \mu_{f(X_1,...,X_n)}(u)$. So, RHS≤LHS. We have already proved that RHS≤LHS; hence, RHS=LHS. Q.E.D.

Proof of Proposition 6. For every α,

$$|X(\alpha)| = \sum_{u \in X(\alpha)} 1.$$

Therefore,

$$\frac{1}{d+2} \sum_{\alpha \in A} |X(\alpha)| = \sum_{u \in U} \left(\frac{1}{d+2} \cdot \sum_{\alpha \in A \& u \in X(\alpha)} 1\right).$$

For every i, if $\mu_X(u) = \alpha_i = (i+1)/(d+2)$, then we have exactly $i+1$ values $(\alpha_0, \alpha_1, ..., \alpha_i)$ for which $u \in X(\alpha)$. Therefore,

$$\sum_{\alpha \in A \& u \in X(\alpha)} 1 = i+1,$$

$$\frac{1}{d+2} \cdot (\sum_{\alpha \in A \& u \in X(\alpha)} 1) = \frac{i+1}{d+2} = \mu_X(u),$$

and so,

$$\frac{1}{d+2} \sum_{\alpha \in A} |X(\alpha)| = \sum_{u \in U} (\frac{1}{d+2} \cdot \sum_{\alpha \in A \& u \in X(\alpha)} 1) = \sum_{u \in U} \mu_X(u).$$

Q.E.D.

Proof of Proposition 7. Since s is a choice function, we have $s([0,1]) \in [0,1]$. Let us denote this number $s([0,1])$ by β. Let us prove that for every interval $[a^-, a^+]$, $s([a^-, a^+]) = \beta \cdot a^+ + (1-\beta)a^-$. We will consider two cases: when the interval is degenerate, and when it is not.

- For a degenerate interval $[a, a]$, from $s([a,a]) \in [a,a]$, we can conclude that $s([a,a]) = a$. But for these intervals, $\alpha a^+ + (1-\beta)a^- = \beta \cdot a + (1-\beta)a = a$, so the formula is correct.
- To get $s([a^-, a^+])$ for an arbitrary non-degenerate interval $[a^-, a^+]$, we can use shift- and unit-invariance:
  - First, we use shift-invariance to prove that
    $$s([a^-, a^+]) = s([(a^+ - a^-) + a^-, 0 + a^-]) = s([a^+ - a^-, 0]) + a^-.$$
  - Then, we use unit-invariance to compute
    $$s([a^+ - a^-, 0]) = s([(a^+ - a^-) \cdot 1, (a^+ - a^-) \cdot 0]) =$$
    $$(a^+ - a^-) s([0,1]) = \beta(a^+ - a^-).$$
  - Finally, we combine these two formulas:
    $$s([a^-, a^+]) = s([a^+ - a^-, 0]) + a^- = \beta(a^+ - a^-) + a^- =$$
    $$\beta a^+ + (1-\beta)a^-.$$

Finally, from symmetry, we conclude that $\beta = 1/2$. Q.E.D.

Proof of Proposition 8 is similar to the proof of Proposition 6.

Acknowledgements

This work was partially supported by NSF Grants No. CDA-9015006 and EEC-9322370, and by NASA Grant No. NAG 9-757. The authors are thankful to R. Baker Kearfott and to the anonymous referees for their help.

REFERENCES

[1] O. Artbauer, "Application of interval, statistical, and fuzzy methods to the evaluation of measurements", *Metrologia*, 1988, Vol. 25, pp. 81–86.

[2] K. T. Atanassov, *Review and new results on intuitionistic fuzzy sets*, Preprint IM-MFAIS-1-88, Sofia, Bulgaria, 1988.

[3] K. T. Atanassov, "Operations over interval valued intuitionistic fuzzy sets", *Fuzzy Sets and Systems*, 1994, Vol. 64, pp. 159–174.

[4] K. T. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets", *Fuzzy Sets and Systems*, 1989, Vol. 31, No. 3, pp. 343–349.

[5] R. E. Bellman, L.A. Zadeh, "Decision making in a fuzzy environment", *Management Science*, 1970, Vol. 17, No. 4, pp. B 141–B 164.

[6] D. Berleant, R. R. Goforth, and A. Khanna, "Comparisons of simulation with pdf's vs. fuzzy values", *Proceedings of the Symposium on Computer Simulation in Industrial Engineering and in Problems of Urban Development*, Nov. 18–19, 1992, Mexico City.

[7] A. K. Bit, M. P. Biswal, S. S. Alam, "Fuzzy programming approach to multicriteria decision making transportation problem", *Fuzzy Sets and Systems*, 1992, Vol. 50, pp. 135–141.

[8] J. M. Blin, "Fuzzy relations in group decision theory", *J. of Cybernetics*, **4**, 17–22, 1974.

[9] J. M. Blin and A. B. Whinston. "Fuzzy sets and social choice", *J. of Cybernetics*, **3**, 28–36, 1973.

[10] G. Bojadziev and G. Bojadzieva, "Liapunov design avoidance control for an open-end two-link robot manipulator", In: M. Jamshidi and M. Saif (eds.), *Robotics and Manufacturing*, ASME Press, N.Y., 1990, Vol. 3, pp. 307–314.

[11] G. Bojadziev and G. Bojadzieva, "Avoidance control of an open-end robot manipulator with two links subject to uncertainties of fuzzy nature", In: Mohammad Jamshidi, Charles Nguyen, Ronald Lumia, and Junku Yuh (Editors), *Intelligent Automation and Soft Computing. Trends in Research, Development, and Applications. Proceedings of the First World Automation Congress (WAC'94), August 14–17, 1994, Maui, Hawaii*, TSI Press, Albuquerque, NM, 1994, Vol. 1, pp. 439–444.

[12] E. L. Deporter, K. P. Ellis, "Optimization of project networks with goal programming and fuzzy linear programming", *Computers in Industrial Engineering*, 1990, Vol. 19, pp. 500–504.

[13] W. M. Dong, W. L. Chiang, H. C. Shah, "Fuzzy information processing in seismic hazard analysis and decision making", *International Journal of Soil Dynamics and Earthquake Engineering*, 1987, Vol. 6, No. 4., pp. 220–226.

[14] W. Dong and F. Wong, "Fuzzy weighted averages and implementation of the extension principle", *Fuzzy Sets and Systems*, 1987, Vol. 21, pp. 183–199.

[15] D. Dubois, M. Grabisch, and H. Prade, "Gradual rules and the approximation of control laws", In: H. T. Nguyen, M. Sugeno, R. Tong, and R. Yager (eds.), *Theoretical aspects of fuzzy control*, J. Wiley, N.Y., 1995, pp. 117–146.

[16] D. Dubois and H. Prade. *Fuzzy sets and systems: theory and applications,* Academic Press, N.Y., London, 1980.

[17] D. Dubois, H. Prade, "Random sets and fuzzy interval analysis", *Fuzzy Sets and Systems*, 1991, Vol. 42, pp. 87–101.

[18] R. Fuller, T. Keresztfalvi, "On generalization of Nguyen's theorem", *Fuzzy Sets and Systems*, 1990, Vol. 4, pp. 371–374.

[19] E. Gardeñes, A. Trepat, J. M. Janer, "Approaches to simulation and to the linear problem in SIGLA system", *Freiburger Intervall-Berichte*, 1981, No. 81/8.

[20] Y. Gentilhomme, "Les ensembles flou en linsguistique", *Cahiers de Ling. Théor. et Appl.*, 1968, Vol. 5, pp. 47–65.

[21] V. A. Gerasimov, M. Yu. Shustrov, "Numerical operations with fuzzy objects", In: S. P. Shary and Yu. I. Shokin (editors), *Interval Analysis*, Krasnoyarsk, Academy of Sciences Computing Center, Technical Report No. 17, 1990, pp. 11–15 (in Russian).

[22] J. A. Goguen, "$L-$fuzzy sets", *Journal of Mathematical Analysis and Applications*, 1967, Vol. 18, pp. 145–174.

[23] J. A. Goguen, "The logic of inexact reasoning", *Synthese*, 1969, **19**, 325–373, 1969 (reprinted in D. Dubois, H. Prade, R. Yager (eds.), *Reading in Fuzzy Sets for Intelligent Systems*, Morgan Kaufmann, San Mateo, CA, 1994, pp. 417–441).

[24] I. R. Goodman, "Algebraic and probabilistic bases for fuzzy sets and the development of fuzzy conditioning", In: I. R. Goodman, M. M. Gupta, H. T. Nguyen, and G. S. Rogers (eds.), *Conditional logic in expert systems*, North Holland, Amsterdam, 1991.

[25] I. R. Goodman, M. M. Gupta, H. T. Nguyen, and G. S. Rogers (eds.), *Conditional logic in expert systems*, North Holland, Amsterdam, 1991.

[26] M. Grabisch, H. T. Nguyen, and E. A. Walker, *Fundamentals of uncertainty calculi with applications to fuzzy inference*, Kluwer, Dordrecht, Netherlands, 1995.

[27] Z. X. Gu and H. Q. Yang, *Basic theory of fuzzy information process*, Chengdu Institute of Radio Engineering Publishing House, P.R. China, 1989 (in Chinese).

[28] A. Kandel, G. Langholtz (editors), *Fuzzy Control Systems*, CRC Press, Boca Raton, FL, 1994.

[29] A. Kaufmann and M. M. Gupta, *Introduction to fuzzy arithmetic*, Van Nostrand Reinhold Co., N.Y., 1985.

[30] R. B. Kearfott, "A Review of Techniques in the Verified Solution of Constrained Global Optimization Problems," *This Volume.*

[31] E. P. Klement, B. Kovalerchuk, "Interval mode of fuzzy control", In: *Proceedings of the International Conference on Interval and Stochastic Methods in Science and Engineering INTERVAL'92*, Moscow, 1992, Vol. 2, pp. 45–46.

[32] G. J. Klir and T. A. Folger, *Fuzzy sets, uncertainty, and information*, Prentice-Hall, U.K., 1988.

[33] G. J. Klir and Bo Yuan, *Fuzzy sets and fuzzy logic. Theory and applications*, Prentice Hall, NJ, 1995.

[34] L. J. Kohout, "Fuzzy interval-valued inference system with para-consistent and grey set extensions", *Reliable Computing*, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, Feb. 23–25, 1995), pp. 107–110.

[35] R. Kruse, J. Gebhardt, and F. Klawonn, *Foundations of fuzzy systems*, J. Wiley & Sons, Chichester, England, 1994.

[36] Y.-J. Lai, C.-L. Hwang, *Fuzzy mathematical programming. Method and application*, Springer-Verlag, Berlin, 1992.

[37] W. A. Lodwick, "Analysis of structure in fuzzy linear programs", *Fuzzy Sets and Systems*, 1990, Vol. 38.

[38] S. G. Loo, *Cybernetica*, 1977, Vol. 20, pp. 201–210.

[39] R. P. Loui, "Interval-based decisions for reasoning systems", In: L. N. Kanal and J. F. Lemmer, *Uncertainty in Artificial Intelligence*, Elsevier, North Holland, 1986, pp. 459–472.

[40] E. H. Mamdani, "Application of fuzzy algorithms for control of simple dynamic plant", *Proceedings of the IEE*, 1974, Vol. 121, No. 12, pp. 1585–1588.

[41] E. H. Mamdani, "Application of fuzzy logic to approximate reasoning using linguistic systems", *IEEE Transactions on Computing*, 1977, Vol. 26, pp. 1182–1191.

[42] R. Lopez De Mantaras, "From intervals to possibility distributions: adding flexibility to reasoning under uncertainty", *Reliable Computing*, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, Feb. 23–25, 1995), pp. 151–152.

[43] M. Mareŝ, *Computation over fuzzy quantities*, CRC Pres, Boca Raton, FL, 1994.

[44] D. Mitra, M. L. Gerard, P. Srinivasan, and A. E. Hands, "A possibilistics interval constraint problem: fuzzy temporal reasoning", *FUZZ-IEEE'94*, IEEE Press, 1994, Vol. 2, pp. 1434–1439.

[45] E. A. Musaev, "The support of interval computations in high-level languages", In: *Proceedings of the 1st Soviet-Bulgarian Seminar on Numerical Processing, Pereslavl-Zalessky, October 19–24, 1987,* Pereslavl-Zalessky, VINITI, 1988, Publ. No. 2634-B89, pp. 110–121 (in Russian).

[46] E. A. Musaev, "Unexpected aspect of interval approach in a task of optimal currency exchange rates", In: *International Conference on Interval and Computer-Algebraic Methods in Science and Engineering (Interval'94), St. Petersburg, Russia, March 7-10, 1994, Abstracts*, pp. 179–180.

[47] A. S. Narin'yani, *Subdefinite sets - a new datatype for knowledge representation*, Academy of Sciences, Siberian Branch, Computing Center, Novosibirsk, Technical Report No. 232, 1980 (in Russian).

[48] A. S. Narin'yani, "Tools that simulate data incompleteness, and their usage in knowledge representation", In: *Knowledge representation and simulation of the understanding process*, Academy of Sciences, Siberian Branch, Computing Center, Novosibirsk, 1980 (in Russian).

[49] A. S. Narin'yani, "Subdefiniteness in knowledge representation and processing systems", *Transactions of USSR Acad. of Sciences, Technical Cybernetics*, 1986, No. 5, pp. 3–28 (in Russian).

[50] H. Nazaraki and I. B. Türksen, "An integrated approach for syllogistic reasoning and knowledge consistency level maintenance", *IEEE Transactions on Systems, Man, and Cybernetics*, 1994, Vol. 24, No. 4, pp. 548–563.

[51] C. V. Negoita and D. A. Ralescu, "Representation theorems for fuzzy concepts", Kybernetes, 1975, Vol. 4, pp. 169–174.

[52] C. V. Negoita and D. A. Ralescu, *Applications of fuzzy sets to systems analysis*, John Wiley and Sons, N.Y., Toronto, 1975.

[53] H. T. Nguyen, "A note on the extension principle for fuzzy sets", *J. Math. Anal. and Appl.*, 1978, Vol. 64, pp. 359–380.

[54] H. T. Nguyen and V. Kreinovich, "Interval Sessions at NAFIPS/IFIS/NASA'94," *Reliable Computing*, 1995, Vol. 1, No. 1, pp. 93–98.

[55] H. T. Nguyen, V. Kreinovich, and Qiang Zuo, "Interval-Valued Degrees of Belief: Applications of Interval Computations to Expert Systems and Intelligent Control" (submitted to *Reliable Computing*).

[56] S. A. Orlovsky, "On programming with fuzzy constraint sets", *Kybernetes*, 1977, Vol. 6, No. 3, pp. 197–201.

[57] S. A. Orlovsky, "Problems of decision-making in case of uncertainty of initial information", Moscow, Nauka Publ., 1981 (in Russian).

[58] M. Sasaki, M. Gen, K. Ida, "A method for solving reliability optimization problems by fuzzy multiobjective 0-1-linear programming", *Electronics and Communications*, 1991, Part 3, Vol. 74, pp. 106–116 (in Japanese).

[59] K. J. Schmucker, *Fuzzy sets, natural language computations, and risk analysis*, Computer Science Press, Rockville, MD, 1984.

[60] Tong Shaocheng, "Interval number and fuzzy number linear programmings", *Fuzzy Sets and Systems*, 1994, Vol. 66, pp. 301–306.

[61] G. Sommer, M. A. Pollatschek, "A fuzzy programming approach to an air pollution regulation problem", *European Journal of Operations Research*, 1978, Vol. 10, pp. 303–313.

[62] N. Tamura and K. Horiuchi, "VSOP fuzzy numbers and fuzzy comparison relations", *Proceedings of the Second IEEE International Conference on Fuzzy Systems*, San Francisco, CA, March 28–April 1, 1993, Vol. II, pp. 1287–1292.

[63] H. Tanaka, T. Okuda, K. Asai, "On fuzzy mathematical programming", *J. of Cybernetics*, 1973, Vol. 3, No. 4, pp. 37–46.

[64] H. Tanaka, K. Asai, "Fuzzy linear programming problems with fuzzy numbers", *Fuzzy Sets and Systems*, 1984, Vol. 13, pp. 1–10.

[65] C.-C. Tsai, C.-H. Chu, T. A. Barta, "Fuzzy linear programming approach to manufacturing cell formation", *FUZZ-IEEE'94*, IEEE Press, 1994, Vol. 2, pp. 1406–1411.

[66] K. Uehara, M. Fujise. "Fuzzy inference based on families of $\alpha-$level sets", *IEEE Transactions on Fuzzy Systems*, 1993, Vol. 1, No. 2, pp. 111–124.

[67] G. Wiedy, H. J. Zimmermann. "Media selection and fuzzy linear programming", *J. Oper. Res. Soc.*, 1978, Vol. 31, pp. 342–249.

[68] R. R. Yager, "A characterization of the extension principle", *Fuzzy Sets and Systems*, 1986, Vol. 18, No. 3, pp. 205–217.

[69] H. Q. Yang, H. Yao, and J. D. Jones, "Calculating functions of fuzzy numbers", *Fuzzy Sets and Systems*, 1993, Vol. 55, pp. 273-283.

[70] A. V. Yazenin, "Fuzzy variables and fuzzy mathematical programming", In: *Models of alternative selection in a fuzzy medium*, Riga Polytechnical Institute, Riga, Latvia, 1984 (in Russian).

[71] A. V. Yazenin, *Fuzzy mathematical programming*, Kalinin University, Kalinin, Russia, 1986 (in Russian).

[72] A. V. Yazenin, "Fuzzy and stochastic programming", *Fuzzy Sets and Systems*, 1987, Vol. 22.

[73] A. V. Yazenin, "Possibility programming models in fuzzy optimization", *Tekhnischeskaya Kibernetika*, 1991, Vol. 29, No. 5 (in Russian); English translation in *Journal of Computer and Systems Sciences International*, 1992, Vol. 31, No. 2.

[74] A. V. Yazenin, "Possibilistic and interval linear programming", *Tekhnischeskaya Kibernetika*, 1993, No. 5, pp. 149–155 (in Russian); English translation in *Journal of Computer and Systems Sciences International*, 1994, Vol. 32, No. 6, pp. 154–160.

[75] J. Yen and N. Pfluger, "Path planning and execution using fuzzy logic", In: *AIAA Guidance, Navigation and Control Conference*, New Orleans, LA, 1991, Vol. 3, pp. 1691–1698.

[76] J. Yen and N. Pfluger, "Designing an adaptive path execution system", In: *IEEE International Conference on Systems, Man and Cybernetics*, Charlottesville, VA, 1991.

[77] J. Yen, N. Pfluger, and R. Langari, "A defuzzification strategy for a fuzzy logic controller employing prohibitive information in command formulation", *Proceedings of IEEE International Conference on Fuzzy Systems, San Diego, CA, March 1992.*

[78] L. Zadeh, "Fuzzy sets", *Information and control*, 1965, Vol. 8, pp. 338–353.

[79] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes", *IEEE Transactions on Systems, Man and Cybernetics*, 1973, Vol. 3, pp. 28–44.

[80] L. A. Zadeh, "The concept of linguistic variable and its application to approximate reasoning", *Information Sciences*, 1975, Vol. 8, pp. 199–249.

[81] L. A. Zadeh, "A theory of common sense knowledge", In: H. J. Scala, S. Termini, and E. Trillas (eds.), *Issues of Vagueness*, Dordrecht, Reidel, 1984, pp. 257–296.

[82] L. A. Zadeh, "Syllogistic reasoning in fuzzy logic and its application to usuality and reasoning with dispositions", *IEEE Trans. SMC,*, 1985, Vol. SMC-15, pp. 754–763.

[83] H. J. Zimmermann, "Description and optimization of fuzzy systems", *Intl. J. of General Systems*, 1976, Vol. 2, No. 4, pp. 209–215.

[84] H. J. Zimmermann, "Fuzzy programming and linear programming with several objective functions", *Fuzzy Sets and Systems*, 1978, Vol. 3, pp. 45–55.

[85] H. J. Zimmermann, "Fuzzy mathematical programming", *Computers and Operations Research*, 1983, Vol. 10, pp. 291–298.

[86] H. J. Zimmermann, *Fuzzy set theory and its applications*, Kluwer, Dordrecht, 1985.

[87] H. J. Zimmermann, "Application of fuzzy set theory to mathematical programming", *Information Sciences*, 1985, Vol. 36, pp. 29–58.

[88] H. J. Zimmermann, *Fuzzy set theory and its applications*, Kluwer, 1991.

[89] V. S. Zyuzin, "An iterative method for solving a system of segment algebraic equations", In: *Differential equations and functions theory*, Saratov University Publ., Saratov, 1987, pp. 72–82 (in Russian).

[90] V. S. Zyuzin, "Twins and a method for solving systems of twin equations", In: *Interval Analysis*, Krasnoyarsk, Academy of Sciences Computing Center, Technical Report No. 6, 1988, pp. 19–21 (in Russian).