1-1-2023

Why Fractional Fuzzy

Mehran Mazandarani
Shenzhen University, memazandarani@szu.edu.cn

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-23-02a

Why Fractional Fuzzy*

Mehran Mazandarani[1] [0000−0003−3116−3841], Olga Kosheleva[2] [0000−0003−2587−4209], and Vladik Kreinovich[2] [0000−0002−1244−1650]

[1] Department of Mechatronics and Control Engineering, Shenzhen University, Shenzhen, China, memazandarani@szu.edu.cn

[2] University of Texas at El Paso, El Paso, Texas, USA, olgak@utep.edu, vladik@utep.edu

Abstract. In many practical situation, control experts can only formulate their experience by using imprecise ("fuzzy") words from natural language. To incorporate this knowledge in automatic controllers, Lotfi Zadeh came up with a methodology that translate the informal expert statements into a precise control strategy. This methodology – and its following modifications – is known as fuzzy control. Fuzzy control often leads to a reasonable control – and we can get an even better control result by tuning the resulting control strategy on the actual system. There are many parameters that can be changed during tuning, so tuning usually is rather time-consuming. Recently, it was empirically shown that in many cases, quite good results can be attained by using a special 1-parametric tuning procedure called fractional fuzzy inference – we get up to 40% improvements just by selecting the proper value of a single parameter. In this paper, we provide a theoretical explanation of why fractional fuzzy inference works so well.

Keywords: Fuzzy control · Fractional fuzzy inference · Tuning.

1 Formulation of the Problem

Need for expert knowledge in control. In some cases – e.g., in controlling a spaceship – we know the exact equations describing the spaceship's trajectory, we know how exactly the spaceship will react to different controls. In such cases, selection of a proper control becomes a mathematical problem.

However, there are also many control situations when an exact model is not known. Such situations are typical in many areas, e.g., in chemical engineering, in medicine, etc. In many such situations, the control is implemented by experts.

* This work was supported in part by the project of the Guoqiang Research Institute of Tsinghua University (No. 2020GQG1001), by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology. It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

The problem is that experts differ in their experience, and there are usually very few top experts, not enough to cover all possible control applications. It is therefore desirable to incorporate the knowledge of top experts into an automatic system that would help others share the benefit of the top expert's knowledge.

Need for fuzzy techniques. Most experts are willing to share their expertise: most of them actually teach students and others. The problem is that they cannot formulate their knowledge in precise numerical terms. This makes perfect sense; e.g., in the US, the vast majority of people can drive, but hardly anyone can answer a question that would naturally arise in automatic control: if you are on a freeway at 100 km/hr, and the car 10 meters in front of you slows down to 95 km/hr, with how many kiloNewtons of force and for how many milliseconds should you press the brake pedal? A natural answer that most driver will give is "press a little bit, for a short time". Such answers – expressed by using imprecise ("fuzzy") words from natural language – are rather typical.

So, to incorporate expert knowledge into a precise control strategy, we need to translate such imprecise statements into precise terms. Techniques for such a translation – pioneered by Lofti Zadeh – are known as *fuzzy techniques*, see, e.g., [1, 4, 7–9, 12].

Need for tuning. In many practical situations, fuzzy techniques provide a reasonable control strategy. However, the resulting control – based on approximate imprecise expert rules – is usually not optimal. To improve the quality of the resulting control, it is necessary to apply it to a real-life system and to "tune it" – i.e., to modify the control strategy based on the results of this application.

Fractional fuzzy techniques are surprisingly successful. A control strategy is a function that assign, to each possible state of the system, an appropriate value(s) of the control. To uniquely determine a function, we need to describe infinitely many numerical values – e.g., the values of this function at all rational inputs. Not surprisingly, most currently used tuning methods tune the values of a large number of parameters – parameters of the corresponding membership functions, etc. (see detailed explanation in the following text). Because we need to determine the values of many different parameters, tuning usually requires a significant amount of computation time.

Interestingly, recently a new tuning technique has been developed – called *fractional fuzzy* technique – that allows to drastically improve the quality of the resulting control by tuning the value of only one parameter; see, e.g., [5, 6]. For example, for the inverted pendulum, this simple 1-parametric tuning leads to a 40% improvement in control quality.

Remaining challenge and what we do in this paper. While fractional fuzzy technique has been empirically successful, there has been no convincing theoretical explanation for its success. In this paper, we provide such an explanation.

The structure of this paper is as follows. To make our explanations clear, in Section 2, we briefly recall how fuzzy techniques work. In Section 3, we describe our main idea – the use of natural invariance, and we show that invariance requirement indeed leads to a few-parametric family that includes fractional fuzzy techniques as a particular case. In Section 4, we show that techniques from this family are actually optimal – in some reasonable sense.

2 Fuzzy Control Techniques: A Brief Reminder

How experts present their knowledge. We want to describe, for each state x – characterized by the values $x_1, \ldots, x_n$ of the corresponding parameters – the appropriate value of the control y. We want to extract such a strategy from the expert statements, and these statements are usually formulated by if-then rules:

if $x_1, \ldots, x_n$ have certain property, then some restrictions are placed on y.

By definition, fuzzy techniques transform expert knowledge into a precise control strategy. So, to describe fuzzy techniques, it is important to recall how experts present their knowledge. This knowledge is usually represented by if-then rules. The most typical situation is when both the conditions and the conclusions of the rules are described by imprecise natural language terms, i.e., when all the rules have the form

if x_1 is A_1 and … and x_n is A_n, then y is B,

where A_i and B are the corresponding terms. For example, we can have a rule

if x_1 is small positive, then y is small negative.

In some cases, experts have a more detailed approximate description of the conclusion, i.e., use rules of the following type

if x_1 is A_1 and … and x_n is A_n, then y is approximately equal to $f(x_1, \ldots, x_n)$,

for some function $f(x_1, \ldots, x_n)$.

What needs to be done to transform this knowledge into a precise control strategy. The expert rules are formed by using imprecise natural-language terms by applying logical connectives like "and" and "if-then". Thus, to transform the experts' if-then rules into a precise control strategy, we need:

- first, to describe natural-language terms like "small" in precise terms, and
- second, to describe how logical connectives – that are usually applied to precise statements – can be applied to the resulting imprecise statements.

Let us describe these two stages one by one.

How to describe natural-language terms like "small" in precise terms. In the original fuzzy technique, to describe an imprecise property A, we assign, to each real number x, the degree (from the interval $[0, 1]$) to which the value x has this property – for example, to which x is small. Here:

– the degree 1 means that we are absolutely sure that x has this property,
– the degree 0 means that we are absolutely sure that x does not have this property, and
– values between 0 and 1 correspond to intermediate degrees of belief.

For most properties, as the input x increases, the degree first (non-strictly) increases, then (non-strictly) decreases. Such properties are known as *fuzzy numbers.*

Where do we get these degrees from? For some values x, we can ask the expert to provide such degrees by marking a number on the scale from 0 to 1. However, there are infinitely many real numbers, and we can only ask finitely many questions. Thus, in practice, we ask the expert about several values, and then use some extrapolation/interpolation techniques to estimate the other degrees. The resulting function assigning a degree $d(x)$ to each real number x is called a *membership function*, or, alternatively, a *fuzzy set.*

For example:

– if we know that the degree $d(x)$ is equal to 0 for $x = x_-$, to 1 for $x = m$, and to 0 for $x = x_+$, for some $x_- < m < x_+$, then linear interpolation leads to a so-called *triangular* membership function;
– if $d(x_-) = d(x_+) = 0$ and $d(m_-) = d(m_+) = 1$ for some $x_- < m_- < m_+ < x_+$, then linear interpolation leads to a so-called *trapezoid* membership function.

For continuous fuzzy numbers, and for each degree $\alpha > 0$, the set of all the values x for which $\mu(x) \geq \alpha$ is an interval. This interval is known as an *α-cut* of the original fuzzy set. Alpha-cuts are nested: if $\alpha < \alpha'$, then the α-cut corresponding to α' is a subset of the α-cut corresponding to α. Once we know all α-cuts $\mathbf{x}(\alpha)$, we can uniquely reconstruct the original membership function as $d(x) = \sup\{\alpha : x \in \mathbf{x}(\alpha)\}$. Thus, the nested family of α-cuts provides an alternative representation of the fuzzy set. This representation is useful in many applications – since it often makes computations easier.

How to describe logical connectives and what to do after that. In situations when each statement is either true or false, the truth value of each composite statement like $A \,\&\, B$ is uniquely determined by the truth values of the component statements A and B. In our case, we only have degrees of confidence in statements A and B, and this information does not uniquely determine the expert's degree of confidence in $A \,\&\, B$.

In the ideal world, we should ask the expert about all such computations. However, in practice, there are too many such combinations, and it is not possible to ask the expect about all of them. It is therefore necessary to be able to estimate the degree of confidence in a combination like $A \,\&\, B$ based only on the available information, i.e., only on the experts' degrees of certainty a and b in the statements A and B. The function that assigns, to each pair of numbers a and b, the corresponding degree is called an *"and"-operation*, or, for historical reasons, a *t-norm.* We will denote the value of the t-norm by $f_\&(a, b)$.

Similarly, if all we know are the degrees of confidence a and b in statements A and B, then our estimate for the degree of confidence in a statement $A \vee B$ will be denoted by $f_\vee(a, b)$, and our estimate for the degree of confidence in an implication $A \to B$ will be denoted by $f_\to(a, b)$.

How can we use these operations? For the general case, when we have rules with imprecise conclusions

if x_1 is A_{k1} and ... and x_n is A_{kn}, then y is B_k,

for $k = 1, \ldots, K$, there are two approaches: logical and Mamdani.

In the logical approach, we estimate the degree of belief $d_k(x_1, \ldots, x_n, y)$ that the k-th rule is satisfied as

$$d_k(x_1, \ldots, x_n, y) = f_\to(f_\&(A_{k1}(x_1), \ldots, A_{kn}(x_n)), B_k(y)),$$

and then compute the degree of belief $d(x_1, \ldots, x_n, y)$ that y is a reasonable control for given data x_i as

$$d(x_1, \ldots, x_n, y) = f_\&(d_1(x_1, \ldots, x_n, y), \ldots, d_K(x_1, \ldots, x_n, y)).$$

In the Mamdani approach, we take into account that y is reasonable if one of the rules is applicable, i.e., if for one of the rules, all conditions are satisfied, and the conclusion is satisfied too. In this case, the degree of belief $d_k(x_1, \ldots, x_n, y)$ that the k-th rule is satisfied is equal to

$$d_k(x_1, \ldots, x_n, y) = f_\&(A_{k1}(x_1), \ldots, A_{kn}(x_n), B_k(y)),$$

and the degree of belief $d(x_1, \ldots, x_n, y)$ that y is a reasonable control for given data x_i is equal to

$$d(x_1, \ldots, x_n, y) = f_\vee(d_1(x_1, \ldots, x_n, y), \ldots, d_K(x_1, \ldots, x_n, y)).$$

In both cases, for each input $x_1, \ldots, x_n$, we get a membership function $m(y) \stackrel{\text{def}}{=} d(x_1, \ldots, x_n, y)$ that describes to what extent different values y are possible. For automatic control, we need to select a single control value $\overline{y}$. The procedure of transforming a (fuzzy) membership function into a single value is known as *defuzzficication*. One of the most widely used defuzzification methods is *centroid defuzzification*, where

$$\overline{y} = \frac{\int y \cdot m(y)\, dy}{\int m(y)\, dy}.$$

In situations when we know an exact description of the conclusion, i.e., when we have rules of the type

if x_1 is A_{k1} and ... and x_n is A_{kn}, then y is approximately equal to $f_k(x_1, \ldots, x_n)$, we first compute the degree $d_k(x_1,\ldots,x_n)$ to which the conditions of each rule are satisfied:

$$d_k(x_1,\ldots,x_n)=f_\&(A_{k1}(x_1),\ldots,A_{kn}(x_n)),$$

and then generate the following control value:

$$\overline{y}=\frac{\sum\limits_{k=1}^{K} d_k(x_1,\ldots,x_n)\cdot f_k(x_1,\ldots,x_n)}{\sum\limits_{k=1}^{K} d_k(x_1,\ldots,x_n)}.$$

Need for tuning. In the above description, we only took into account the expert's imprecise knowledge. To get a more adequate control, we need to test it on a real-life system, and make adjustments if needed. This real-system-based procedure is known as *tuning*.

3 Fractional Fuzzy Techniques: Motivations, Description, Successes, and Remaining Challenge

Motivations: need for faster tuning techniques. In general, there are many parameters to tune: e.g., the parameters describing all the membership functions $A_{ki}(x_i)$ and $B_k(y)$. Such tuning takes a lot of computation time. To speed up computations, it is therefore desirable to come up with tuning methods that require only a small number of parameters.

Fractional fuzzy techniques: description. Recently, a new few-parametric tuning method was proposed; see, e.g., [5, 6]. There are, three versions of this method:

- In the first version, we select a real number $\beta_+>0$, and we replace each α-cut interval $[\underline{x},\overline{x}]$ with a new interval $[\underline{x},\underline{x}+\beta_+\cdot(\overline{x}-\underline{x})]$.
- In the second version, we select a real number $\beta_-<1$, and we replace each α-cut interval $[\underline{x},\overline{x}]$ with a new interval $[\underline{x}+\beta_-\cdot(\overline{x}-\underline{x}),\overline{x}]$.
- In the combined third version, we select two numbers $\beta_-\le\beta_+$, and we replace each α-cut interval $[\underline{x},\overline{x}]$ with a new interval

$$[\underline{x}+\beta_-\cdot(\overline{x}-\underline{x}),\underline{x}+\beta_+\cdot(\overline{x}-\underline{x})].$$

In each version of this method, we replace each α-cut with its fraction; thus, this method is known as *fractional fuzzy* technique.

Fractional fuzzy techniques: successes. Practical applications show that these techniques work very well. For example, for the inverted pendulum, each of the first two versions – corresponding to 1-parametric tuning – leads to a 40% improvement in control quality [5, 6].

Fractional fuzzy techniques: remaining challenge, and what we do in this paper. A natural question is: how can we explain this empirical success? In this paper, we explain why this method is so successful.

4 Our Main Idea and How It Explains the Empirical Success of Fractional Fuzzy Techniques

Experts are not perfect. Suppose that the best estimate of the corresponding quantity is $\widetilde{x}$, and the actual uncertainty – that can be derived from what we know – is $\pm\Delta$, meaning that the actual value of the quantity x is somewhere in the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. This is what an ideal expert should return.

Actual experts are not perfect, they produce intervals $[x_0 - \delta, x_0 + \delta]$ which are, in general, different from the ideal interval:

- an expert may overestimate the value of the quantity x, by producing a larger value $x_0 > \widetilde{x}$;
- an expert may underestimate this value, by producing $x_0 < \widetilde{x}$;
- an expert may overestimate the inaccuracy, by producing a value $\delta > \Delta$;
- an expert may underestimate the inaccuracy, by producing a value $\delta < \Delta$.

In all these cases, the interval $[x_0 - \delta, x_0 + \delta]$ produced by an expert is different from the desired interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$.

So, to make a control more adequate, a natural idea is to take this into account and to transform the expert's interval back into the desired interval. For this purpose, we need to come up with a transformation T that transforms intervals into intervals.

Natural properties of the transformation T. Both inputs and outputs of the transformation T are intervals of values of a physical quantity, i.e., intervals for which both endpoints are values of this quantity. We would like to deal with the actual values, but in practice, we can only deal with numerical values, and numerical values depend on what unit we choose for this quantity and what starting point we choose.

If we select a measuring unit which is λ times smaller than the original one, then all the numerical values are multiplied by λ: $x \mapsto \lambda \cdot x$. For example, if we replace meters with centimeters, then 1.7 m becomes $1.7 \cdot 100 = 170$ cm.

If we select a new starting point which is x_0 units smaller than the original one, then this value x_0 is added to all the numerical values: $x \mapsto x + x_0$. For example, if we replace Celsius scale for temperature to Kelvin, then we need to add 273 to all the numerical values.

The choices of a measuring unit and of a starting point are often arbitrary, coming from a reasonably arbitrary agreement. It is therefore reasonable to require that the desired transformation T lead to the same interval of real values. So, we arrive at the following definitions.

Definition 1. *We say that the mapping T from intervals to intervals is* scale-invariant *if for every interval $[a, b]$ and for every real number $\lambda > 0$, the equality $[c, d] = T([a, b])$ implies that $[c', d'] = T([a', b'])$, where we denoted $a' = \lambda \cdot a$, $b' = \lambda \cdot b$, $c' = \lambda \cdot c$, and $d' = \lambda \cdot d$.*

Definition 2. *We say that the mapping T from intervals to intervals is* shift-invariant *if for every interval $[a, b]$ and for every real number x_0, the equality*

$[c,d] = T([a,b])$ implies that $[c',d'] = T([a',b'])$, where we denoted $a' = a + x_0$, $b' = b + x_0$, $c' = c + x_0$, and $d' = d + x_0$.

Comment. Similar invariance conditions were used in [2] to explain another empirically successful interval transformations [3, 10, 11]. However, these papers only dealt with overestimation or underestimation of uncertainty. Our analysis analyzes a more general situation, where the estimate itself can also be biased.

Proposition 1. *A mapping T is scale- and shift-invariant if and only if it has the form*

$$T([a,b]) = [a + \beta_- \cdot (b-a), a + \beta_+ \cdot (b-a)]$$

for some $\beta_- \le \beta_+$.

Comment.

- For $\beta_- = 0$, we get the first version of the fractional fuzzy techniques.
- For $\beta_+ = 1$, we get the second version of the fractional fuzzy techniques.
- In general, we get the third (combined) version of these techniques.

Thus, indeed, this proposition provides an explanation for fractional fuzzy techniques.

Proof. Let us denote the endpoint of the interval $T([0,1])$ by, correspondingly, β_- and β_+, i.e., $T([0,1]) = [\beta_-, \beta_+]$. Let us show that for every interval $[a,b]$, the result $T([a,b])$ of applying this transformation has the desired form.

Indeed, due to scale-invariance for $\lambda = b - a$, we have

$$T([0, b-a]) = [\beta_- \cdot (b-a), \beta_+ \cdot (b-a)].$$

Now, due to shift-invariance with $x_0 = a$, we get the desired formula

$$T([a,b]) = [a + \beta_- \cdot (b-a), a + \beta_+ \cdot (b-a)].$$

The proposition is proven.

5 Fractional Fuzzy Techniques Are Optimal – In Some Reasonable Sense

What do we mean by optimal. Usually, when people talk about optimality, they assume that there is some numerical criterion, and the optimal alternative is the one that has the largest (or the smallest) value of this criterion. For example:

- an optimal path may be the shortest path,
- an optimal investment portfolio is the one with the largest expected gain, etc.

However, this is not the most general description of optimality. For example, if we have two alternative investment portfolios with the same expected gain, it is reasonable to select the one with the smallest expected deviation from this gain, etc. Thus, the optimality criterion can be more complicated than simply comparing numerical values.

In general, what we want from an optimality criterion is that it should allow us, at least for some pairs of alternatives a, a' to decide:

- whether a is better than a' (we will denote it by $a' < a$)
- or a' is better than a ($a < a'$),
- or a and a' are the same quality to the user (we will denote it by $a \sim a'$).

Of course, there should be natural transitivity requirements: e.g., if a is better than a' and a' is better than a'', then a should be better than a''.

As we have mentioned in the beginning of the previous paragraph, if we have several alternatives which are optimal with respect to some optimality criterion, this means that this criterion is not final: we can use this non-uniqueness to optimize something else. So, when the criterion is final, there is exactly one alternative that is optimal with respect to this criterion. Thus, we arrive at the following definitions.

Definition 3. *By an* optimality criterion *on a set A, we mean a pair $(<, \sim)$ of binary relations on this set gthat satisfy the following properties:*

- *if $a < b$ and $b < c$, then $a < c$;*
- *if $a < b$ and $b \sim c$, then $a < c$;*
- *if $a \sim b$ and $b < c$, then $a < c$;*
- *if $a \sim b$ and $b \sim c$, then $a \sim c$;*
- *if $a < b$ then $a \not\sim b$;*
- *if $a \sim b$, then $b \sim a$;*
- *always $a \sim a$.*

Definition 4. *We say that an alternative a is* optimal *with respect to an optimality criterion $(< \,.\, \sim)$ if for every $b \in A$, we have either $b < a$ or $b \sim a$.*

Definition 5. *We say that an optimality criterion $(<, \sim)$ is* final *if there is exactly one alternative that is optimal with respect to this criterion.*

The optimality criterion should be invariant. In our case, alternatives are transformation functions. It is reasonable to require that if one transformation is better than another one, then it will still be better if we use a different measuring unit or a different starting point. Let us describe this in precise terms.

Suppose that we have an interval $[a, b]$ expressed in the original units. If we use a new measuring unit which is λ times smaller, then the interval becomes $\lambda \cdot [a, b] \stackrel{\text{def}}{=} [\lambda \cdot a, \lambda \cdot b]$. If we apply a transformation T to this new interval, we get the interval $T([\lambda \cdot a, \lambda \cdot b])$. This interval is in the new units; in the original units, it will have the form $\lambda^{-1} \cdot T([\lambda \cdot a, \lambda \cdot b])$. This is equivalent to using a transformation T_λ for which

$$T_\lambda([a, b]) = \lambda^{-1} \cdot T([\lambda \cdot a, \lambda \cdot b]). \tag{1}$$

In these terms, invariance means that:

- if $T < T'$, then we should have $T_\lambda < T'_\lambda$, and
- if $T \sim T'$, then we should have $T_\lambda \sim T'_\lambda$.

Similarly, suppose that we have an interval $[a, b]$ corresponding to the original starting point. If we use a new starting point which is x_0 units smaller, then the interval becomes $[a, b] + x_0 \stackrel{\text{def}}{=} [a + x_0, b + x_0]$. If we apply a transformation T to this new interval, we get the interval $T([a + x_0, b + x_0])$. The endpoints of this interval correspond to the new starting point; with respect to the original starting point, it will have the form $T([a + x_0, b + x_0]) - x_0$. This is equivalent to using a transformation $T_{(x_0)}$ for which

$$T_{(x_0)}([a, b]) = T([a + x_0, b + x_0]) - x_0. \tag{2}$$

In these terms, invariance means that:

- if $T < T'$, then we should have $T_{(x_0)} < T'_{(x_0)}$, and
- if $T \sim T'$, then we should have $T_{(x_0)} \sim T'_{(x_0)}$.

Definition 6. *We say that an optimality criterion on the set of all interval-to-interval transformation is* scale-invariant *if for every two transformations T and T' and for every real number $\lambda > 0$,*

- *$T < T'$ implies $T_\lambda < T'_\lambda$ and*
- *$T \sim T'$ implies $T_\lambda \sim T'_\lambda$,*

where T_λ and T'_λ are described by the formula (1).

Definition 7. *We say that an optimality criterion on the set of all interval-to-interval transformations is* shift-invariant *if for every two transformation T and T' and for every real number x_0,*

- *$T < T'$ implies $T_{(x_0)} < T'_{(x_0)}$ and*
- *$T \sim T'$ implies $T_{(x_0)} \sim T'_{(x_0)}$,*

where $T_{(x_0)}$ and $T'_{(x_0)}$ are described by the formula (2).

Proposition 2. *For every scale-invariant, shift-invariant, and final optimality criterion, the optimal transformation has the form*

$$t([a, b]) = [a + \beta_- \cdot (b - a), a + \beta_+ \cdot (b - a)]$$

for some $\beta_- \le \beta_+$.

Comment. The optimal transformation has exactly the form used in fractional fuzzy techniques. Thus, this result provides a theoretical explanation for the empirical fact that these techniques work well in many control situations.

Proof. Let $(<, \sim)$ be scale-invariant, shift-invariant, and final optimality criterion, and let t be optimal with respect to this criterion.

Let us prove that t is scale-invariant, i.e., that for all λ, we have $t_\lambda = t$. Indeed, since t is optimal, for every T, we have $T < t$ or $T \sim t$. In particular, this is true for $T_{\lambda^{-1}}$, i.e., we have $T_{\lambda^{-1}} < t$ or $T_{\lambda^{-1}} \sim t$. Due to the fact that the optimality criterion is scale-invariant, we conclude that for every T, we have either $T < t_\lambda$ or $T \sim t_\lambda$. By definition of optimality, this means that the transformation t_λ is optimal. However, t is also optimal, and we assumed that the optimality criterion is final, i.e., that there is only one optimal alternative. Thus, $t_\lambda = t$.

Similarly, we can prove that the transformation t is shift-invariant. Thus, by Proposition 1, the transformation t has the desired form.

6 Conclusions

In many areas of human activity, there are people who are very good in the corresponding tasks: top medical doctors excel in diagnosing and treating diseases, top pilots excel in piloting planes, etc. It is desirable to incorporate their expertise into automated systems that would help others make similarly effective decisions – or even make these decision by themselves, without the need for a human controller. These top folks are usually willing to share their knowledge and their skills, but the problem is that they often formulate a significant part of their skills not in precise numerical terms, but by using imprecise ("fuzzy") words from natural language, like "small". To transform such knowledge into numerical computer-understandable form, Lotfi Zadeh invented fuzzy techniques. In these techniques, we first translate expert knoweldge into numerical terms, and then tune the resulting control so as to make it as effective as possible.

This procedure has led to many successful applications. However, in many cases, achieving this success required a lot of time and efforts: indeed, there are usually many parameters to tune, and, as a result, tuning is often very time-consuming. To speed up the tuning process, it is desirable to come up with effective few-parametric tuning procedures. In this paper, we analyze one such procedure – known as fractional fuzzy techniques – in which we replace each α-cut interval with its (appropriately defined) fraction. This procedure turned out to be very effective – e.g., it improves the quality of decisions by up to 40% in the case of the reverse pendulum problem.

A natural question is: how to explain this empirical success? In this paper, we provide a theoretical explanation for this success: namely, we show that the corresponding few-parametric family of tunings is, in some reasonable sense, optimal.

Acknowledgements The authors are greatly thankful to all the participants of the 2022 IEEE International Conference on Systems, Man, and Cybernetics

(Prague, Czech Republic, October 9–12, 2022) for valuable discussions, and to the anonymous referees for useful suggestions.

References

1. Belohlavek, R., Dauben, J. W., Klir, G. J.: Fuzzy Logic and Mathematics: A Historical Perspective. Oxford University Press, New York (2017)
2. Berrout Ramos, L. A., Kreinovich, V., Autchariyapanitkul, K.: Correcting interval-valued expert estimates: empirical formulas explained. In: Thach, N. N., Kreinovich, V., Ha, D. T., Trung, N. D. (eds.), Financial Econometrics: Bayesian Analysis, Quantum Uncertainty, and Related Topics. pp. 1–7, Springer, Cham, Switzerland (2022)
3. Gajdos, T., Vergnaud, J.-C.: Decisions with conflicting and imprecise information. Social Choice and Welfare **41**, 427–452 (2013)
4. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic. Prentice Hall, Upper Saddle River, New Jersey (1995)
5. Mazandarani, M., Xiu, L.: Fractional fuzzy inference system: The new generation of fuzzy inference systems. IEEE Access **8**, 126066–126082 (2020)
6. Mazandarani1, M., Xiu, L.: Interval type-2 fractional fuzzy inference systems: Towards an evolution in fuzzy inference systems. Expert Systems With Applications **189**, Paper 115947 (2022)
7. Mendel, J. M.: Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions. Springer, Cham, Switzerland (2017)
8. Nguyen, H. T., Walker, C. L., Walker, E. A.: A First Course in Fuzzy Logic. Chapman and Hall/CRC, Boca Raton, Florida (2019)
9. Novák, V., Perfilieva, I., Močkoř, J.: Mathematical Principles of Fuzzy Logic. Kluwer, Boston, Dordrecht (1999)
10. Smithson, M.: Conflict and ambiguity: preliminary models and empirical tests. In: Proceedings of the Eighth International Symposium on Imprecise Probability: Theories and Applications ISIPTA'2013, Compiegne, France, July 2–5, 2013, pp. 303–310 (2013) '
11. Smithson, M.: Probability judgments under ambiguity and conflict. Frontiers in Psychology **6**, Paper 674 (2015)
12. Zadeh, L. A.: Fuzzy sets. Information and Control **8**, 338–353 (1965)