University of Texas at El Paso

ScholarWorks@UTEP

Departmental Technical Reports (CS)   Computer Science

6-2020

Natural Invariance Explains Empirical Success of Specific Membership Functions, Hedge Operations, and Negation Operations

Julio Urenda
The University of Texas at El Paso, jcurenda@utep.edu

Orsoly Csiszár
Óbuda University, orsolya.csiszar@nik.uni-obuda.hu

Gábor Csiszár
University of Stuttgart, gabor.csiszar@mp.imw.uni-stuttgart.de

József Dombi
University of Szeged, dombi@inf.u-szeged.hu

György Eigner
Óbuda University, eigner.gyorgy@nik.uni-obuda.hu

Follow this and additional works at: https://scholarworks.utep.edu/cs_techrep

Part of the Computer Sciences Commons

Comments:
See next page for additional authors
Technical Report: UTEP-CS-20-58

To appear in: *Proceedings of the Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2020*, Redmond, Washington, August 20-22, 2020.

Authors

Julio Urenda, Orsoly Csiszár, Gábor Csiszár, József Dombi, György Eigner, and Vladik Kreinovich

Natural Invariance Explains Empirical Success of Specific Membership Functions, Hedge Operations, and Negation Operations

Julio C. Urenda[1,2], Orsolya Csiszár[3,4],
Gábor Csiszár[5], József Dombi[6], György Eigner[4], and
Vladik Kreinovich[1]
[1]Department of Computer Science
[2]Department of Mathematical Sciences
University of Texas at El Paso
El Paso, TX 79928, USA
jcurenda@utep.edu, vladik@utep.edu
[3]Faculty of Basic Sciences, University of Applied Sciences Esslingen
Esslingen, Germany
[4]Institute of Applied Mathematics, Óbuda University
Budapest, Hungary, orsolya.csiszar@nik.uni-obuda.hu
eigner.gyorgy@nik.uni-obuda.hu
[5]Institute of Materials Physics, University of Stuttgart
Stuttgart, Germany
gabor.csiszar@mp.imw.uni-stuttgart.de
[6]Institute of Informatics, University of Szeged
Szeged, Hungary, dombi@inf.u-szeged.hu

Abstract

Empirical studies have shown that in many practical problems, out of all symmetric membership functions, special *distending* functions work best, and out of all hedge operations and negation operations, fractional linear ones work the best. In this paper, we show that these empirical successes can be explained by natural invariance requirements.

1   Formulation of the Problem

Fuzzy techniques: a brief reminder. In many applications, we have knowledge formulated in terms of imprecise ("fuzzy") terms from natural language, like "small", "somewhat small", etc. To translate this knowledge into computer-understandable form, Lotfi Zadeh proposes *fuzzy techniques*; see, e.g., [1, 11, 12,

1

14, 15, 17]. According to these techniques, each imprecise property like "small" can be described by assigning, to each value x of the corresponding quantity, a degree $\mu(x)$ to which, according to the expert, this property is true. These degrees are usually selected from the interval $[0, 1]$, so that 1 corresponds to full confidence, 0 to complete lack of confidence, and values between 0 and 1 describe intermediate degrees of confidence. The resulting function $\mu(x)$ is known as a *membership function*.

In practice, we can only ask finitely many questions to the expert, so we only elicit a few values $\mu(x_1)$, $\mu(x_2)$, etc. Based on these values, we need to estimate the values $\mu(x)$ for all other values x. For this purpose, usually, we select a family of membership functions – e.g., triangular, trapezoidal, etc. – and select a function from this family which best fits the known values.

For terms like "somewhat small", "very small", the situation is more complicated. We can add different "hedges" like "somewhat", "very", etc., to each property. As a result, we get a large number of possible terms, and it is not realistically possible to ask the expert about each such term. Instead, practitioners estimate the degree to which, e.g., "somewhat small" is true based on the degree to which "small" is true. In other words, with each linguistic hedge, we associate a function h from $[0, 1]$ to $[0, 1]$ that transform the degree to which a property is true into an estimate for the degree to which the hedged property is true.

Similarly to the membership functions, we can elicit a few values $h(x_i)$ of the hedge operation from the experts, and then we extrapolate and/or interpolate to get all the other values of $h(x)$. Usually, a family of hedge operations is pre-selected, and then we select a specific operation from this family which best fits the elicited values $h(x_i)$.

Similarly, instead of asking experts for their degrees of confidence in statements containing negation, such as "not small", we estimate the expert's degree of confidence in these statements based on their degrees of confidence in the positive statements. The corresponding operation $n(x)$ is known as the *negation operation*.

Need to select proper membership functions, proper hedge operations, and proper negation operations. Fuzzy techniques have been successfully applied to many application areas. However, this does not necessarily mean that every time we try to use fuzzy techniques, we get a success story. The success (or not) often depends on which membership functions and which hedge and negation operations we select: for some selections, we get good results (e.g., good control), for other selections, the results are not so good.

What we do in this paper. There is a lot of empirical data about which selections work better. In this paper, we provide a general explanation for several of these empirically best selections, an explanation based on the natural concepts of invariance.

Specifically, we explain the following empirically successful selections:

- for symmetric membership functions that describe properties like "small", for which $\mu(x) = \mu(-x)$ and the degree $\mu(|x|)$ decreases with $|x|$, in many

practical situations, the most empirically successful are so-called *distend-ing* membership functions, i.e., functions of the type

$$\mu(x) = \frac{1}{1 + a \cdot |x|^b} \qquad (1)$$

for some a and b; see, e.g., [7, 8, 9];

- among hedge and negation operations, in many practical situations, the most efficient are fractional linear functions

$$h(x) = \frac{a + b \cdot x}{1 + c \cdot x} \qquad (2)$$

for some a, b, and c; see, e.g., [2, 3, 4, 5].

2   Analysis of the Problem

Re-scaling. The variable x describes the value of some physical quantity, such a distance, height, difference in temperatures, etc. When we process these values, we deal with numbers, but numbers depend on the selection of the measuring unit: if we replace the original measuring unit with a new one which is λ times smaller, then all the numerical values will be multiplied by λ: $x \to X = \lambda \cdot x$. For example, 2 meters become $2 \cdot 100 = 200$ centimeters. This transformation from one measuring scale to another is known as *re-scaling*.

Scale-invariance: idea. In many physical situations, the choice of a measuring unit is rather arbitrary. In such situations, all the formulas remain the same no matter what unit we use.

For example, the formula $y = x^2$ for the area of the square with side x remains valid if we replace the unit for measuring sides from meters with centimeters – of course, we then need to appropriately change the unit for y, from square meters to square centimeters. In general, invariance of the formula $y = f(x)$ means that for each re-scaling $x \to X = \lambda \cdot x$, there exists an appropriate re-scaling $y \to Y$ for which the same formula $Y = f(X)$ will be true for the correspondingly re-scaled variables X and Y.

Let us apply this idea to the membership function. It is reasonable to require that the selection of the best membership functions should also not depend on the choice of the unit for measuring the corresponding quantity x. In other words, it is reasonable to require that for each $\lambda > 0$, there should exist some reasonable transformation $y \to Y = T(y)$ of the degree of confidence for which $y = \mu(x)$ implies $Y = \mu(X)$.

So, what are reasonable transformations of the degree of confidence? One way to measure the degree of confidence is to have a poll: ask N experts how many of them believe that a given value x is, e.g., small, count the number M of whose who believe in this, and take the ratio M/N as the desired degree $y = \mu(x)$.

As usual with polls, the more people we ask, the more adequately we describe the general opinion. So, to get a more accurate estimate for $\mu(x)$, it is reasonable to ask more people. When we have a limited number of people to ask, it is reasonable to ask top experts in the field. When we start asking more people, we are thus adding people who are less experienced – and who may therefore be somewhat intimidated by the opinions of the top experts. This intimidation can be expressed in different ways:

- some new people may be too shy to express their own opinion, so they will keep quiet; as a result, if we add A people to the original N, we sill still have the same number M of people voting "yes", and the new ratio will be equal to $Y = \dfrac{M}{N + A}$, i.e., to $Y = a \cdot y$, where $a \overset{\text{def}}{=} \dfrac{N}{N + A}$;

- some new people will be too shy to think on their own and will vote with the majority; so for the case when $M > N/2$, we will have

$$Y = \frac{M + A}{N + A},$$

i.e., since $M = y \cdot N$, we will have

$$Y = \frac{y \cdot N + A}{N + A} = a \cdot y + b,$$

where a is the same as before and $b = \dfrac{A}{N + A}$;

- we may also have a situation in which a certain proportion c of the new people keep quiet while the others vote with the majority; in this case, we have
$$Y = \frac{M + (1 - c) \cdot A}{N + A} = a \cdot y + b,$$

where $a = (1 - c) \cdot \dfrac{A}{N + A}$.

In all these cases, we have a linear transformation $Y = a \cdot y + b$. So, it seems reasonable to identify reasonable transformations with linear ones. We will call the corresponding scale-invariance L-scale-invariance (L for Linear).

What membership functions we consider. We consider symmetric properties, for which $\mu(-x) = \mu(x)$, so it is sufficient to consider only positive values x. Specifically, we properties like "small" for which the degree of confidence decreases with x, going all the way to 0 as x increases. We will call such membership functions s-membership functions (s for small). Thus, we arrive at the following definition.

Definition 1. *By an* s-membership function, *we means a function $\mu : (0, \infty) \to [0, 1]$ that, starting with $\mu(0) = 1$, decreases with x (i.e., for which $x_1 > x_2$ implies $\mu(x_1) \geq \mu(x_2)$) and for which $\lim_{x \to \infty} = 0$.*

Definition 2. *We say that an s-membership function $\mu(x)$ is* L-scale-invariant *if for every $\lambda > 0$, there exist values $a(\lambda)$ and $b(\lambda)$ for which $y = \mu(x)$ implies $Y = \mu(X)$, where $X = \lambda \cdot x$ and $Y = a(\lambda) \cdot y + b(\lambda)$.*

Unfortunately, this does not solve our problem: as the following result shows, the only L-scale-invariant s-membership functions are constants;

Proposition 1. *The only L-scale-invariant s-membership functions are constant functions $\mu(x) = \text{const}$.*

Discussion. What does this result mean? We considered two possible types of reasonable transformations of the degrees of confidence – which both turned out to be linear, and this was not enough. So probably there are other reasonable transformations of degrees of confidence. How can we describe such transformations?

Clearly, if we have a reasonable transformation, then its inverse is also reasonable. Also, a composition of two reasonable transformations should be a reasonable transformation too. So, in mathematical terms, reasonable transformations should form a *group*.

This group should be finite-dimensional, in the sense that different transformations should be uniquely determined by a finite number of parameters – since in the computer, we can store only finitely many parameters. We also know that linear transformations are reasonable. So, we are looking for a finite-dimensional group of transformations from real numbers to real numbers that contains all linear transformations. It is known (see, e.g., [10, 13, 16]) that all such transformations are piece-wise linear, i.e., have the form

$$\mu \to \frac{a \cdot \mu + b}{1 + c \cdot \mu}.$$

Thus, we arrive at the following definitions.

3 Which Symmetric Membership Functions Should We Select: Definitions and the Main Result

Definition 3. *We say that an s-membership function $\mu(x)$ is* scale-invariant *if for every $\lambda > 0$, there exist values $a(\lambda)$, $b(\lambda)$, and $c(\lambda)$ for which $y = \mu(x)$ implies $Y = \mu(X)$, where $X = \lambda \cdot x$ and*

$$Y = \frac{a(\lambda) \cdot y + b(\lambda)}{1 + c(\lambda) \cdot y}.$$

Proposition 2. *The only scale-invariant s-membership functions are distending membership functions (1).*

Discussion. This result explains the empirical success of distending functions.

4 Which Hedge Operations and Negation Operations Should We Select

Discussion. We would like hedging and negation operations $y = h(x)$ to be also invariant, i.e., that for each natural transformation $X = T(x)$, there should be a transformation $Y = S(y)$ for which $y = h(x)$ implies $Y = h(X)$. Now that we know what are natural transformations of membership degrees – they are fractional-linear functions – we can describe this requirement in precise terms.

Definition 4. *We say that a monotonic function $y = h(x)$ from an open (finite or infinite) interval D to real numbers is* h-scale-invariant *if for every fractional-linear transformation $X = T(x)$, there exists a fractional-linear transformation $Y = S(y)$ for which $y = h(x)$ implies $Y = h(X)$.*

Proposition 3. *The only h-scale-invariant functions are fractionally linear ones.*

Discussion.

- This result explains the empirical success of fractional-linear hedge operations and negation operations.

- As we show in the proof, it is sufficient to require that a fractional linear transformation S exists only for all *linear* transformations T.

5 Proofs

5.1 Proof of Proposition 1.

We will prove this result by contradiction. Let us assume that the function $\mu(x)$ is not a constant, and let us derive a contradiction.

Substituting the expressions for X, Y, and $y = \mu(x)$ into the formula $Y = \mu(X)$ describing L-scale-invariance, we conclude that for every x and for every λ, we have

$$\mu(\lambda \cdot x) = a(\lambda) \cdot \mu(x) + b(\lambda). \tag{3}$$

It is known that monotonic functions are almost everywhere differentiable. Due to the formula (3), if a function $\mu(x)$ is differentiable at some point $x = x_0$, it is also differentiable at any point of the type $\lambda \cdot x_0$ for every $\lambda > 0$ – and thus, that it is differentiable for all $x > 0$.

Since the function $\mu(x)$ is not constant, there exist values $x_1 \neq x_2$ for which $\mu(x_1) \neq \mu(x_2)$. For these values, the formula (3) has the form

$$\mu(\lambda \cdot x_1) = a(\lambda) \cdot \mu(x_1) + b(\lambda); \quad \mu(\lambda \cdot x_2) = a(\lambda) \cdot \mu(x_2) + b(\lambda).$$

Subtracting the two equations, we get

$$\mu(\lambda \cdot x_1) - \mu(\lambda \cdot x_2) = a(\lambda) \cdot (\mu(x_1) - \mu(x_2)),$$

thus
$$a(\lambda) = \frac{\mu(\lambda \cdot x_1) - \mu(\lambda \cdot x_2)}{\mu(x_1) - \mu(x_2)}.$$

Since the function $\mu(x)$ is differentiable, we can conclude that the function $a(\lambda)$ is also differentiable. Thus, the function $b(\lambda) = \mu(\lambda \cdot x) - a(\lambda) \cdot \mu(x)$ is differentiable too.

Since all three functions $\mu(x)$, $a(\lambda)$, and $b(\lambda)$ are differentiable, we can differentiate both sides of the equality (3) with respect to λ. If we substitute $\lambda = 1$, we get $x \cdot \mu'(x) = A \cdot \mu(x) + B$, where we denoted $A \stackrel{\text{def}}{=} a'(1)$, $B \stackrel{\text{def}}{=} b'(1)$, and $\mu'(x)$, as usual, indicates the derivative. Thus, $x \cdot \dfrac{d\mu}{dx} = A \cdot \mu + B$. We cannot have $A = 0$ and $B = 0$, since then $\mu'(x) = 0$ and $\mu(x)$ would be a constant. Thus, in general, the expression $A \cdot \mu + B$ is not 0, so

$$\frac{d\mu}{A \cdot \mu + B} = \frac{dx}{x}.$$

If $A = 0$, then integration leads to $\dfrac{1}{B} \cdot \mu(x) = \ln(x) + c$, where c_0 is the integration constant. Thus, $\mu(x) = B \cdot \ln(x) + B \cdot c_0$. This expression has negative values for some x, while all the values $\mu(x)$ are in the interval $[0, 1]$. So, this case is impossible.

If $A \neq 0$, then we have $d(A \cdot \mu + B) = A \cdot d\mu$, hence

$$\frac{d(A \cdot \mu + B)}{A \cdot \mu + B} = A \cdot \frac{dx}{x}.$$

Integration leads to $\ln(A \cdot \mu(x) + B) = A \cdot \ln(x) + c_0$. By applying $\exp(z)$ to both sides, we get $A \cdot \mu(x) + B = \exp(c_0) \cdot x^A$, i.e., $\mu(x) = A^{-1} \cdot \exp(c_0) \cdot x^A - B/A$. This expression tends to infinity either for $x \to \infty$ (if $A > 0$) or for $x \to 0$ (if $A < 0$). In both cases, we get a contradiction with our assumption that $\mu(x)$ is always within the interval $[0, 1]$.

The proposition is proven.

5.2 Proof of Proposition 2

Substituting the expressions for X, Y, and $y = \mu(x)$ into the formula $Y = \mu(X)$ describing scale-invariance, we conclude that for every x and for every λ, we have

$$\mu(\lambda \cdot x) = \frac{a(\lambda) \cdot \mu(x) + b(\lambda)}{1 + c(\lambda) \cdot \mu(x)}. \tag{4}$$

Similarly to the proof of Proposition 1, we can conclude that the function $\mu(x)$ is differentiable for all $x > 0$.

Multiplying both sides of the equality (4) by the denominator, we conclude that

$$\mu(\lambda \cdot x) + c(\lambda) \cdot \mu(x) \cdot \mu(\lambda \cdot x) = a(\lambda) \cdot \mu(x) + b(\lambda).$$

7

So, for three different values x_i, we have the following three equations:

$$\mu(\lambda \cdot x_i) + c(\lambda) \cdot \mu(x_i) \cdot \mu(\lambda \cdot x_i) = a(\lambda) \cdot \mu(x_i) + b(\lambda), \quad i = 1, 2, 3.$$

We thus have a system of three linear equations for three unknowns $a(\lambda)$, $b(\lambda)$, and $c(\lambda)$. By Cramer's rule, the solution to such a system is a rational (hence differentiable) function of the coefficients and the right-hand sides. So, since the function $\mu(x)$ is differentiable, we can conclude that the functions $a(\lambda)$, $b(\lambda)$, and $c(\lambda)$ are differentiable as well.

Since all the functions $\mu(x)$, $a(\lambda)$, $b(\lambda)$, and $c(\lambda)$ are differentiable, we can differentiate both sides of the formula (4) with respect to λ. If we substitute $\lambda = 1$ and take into account that for $\lambda = 1$, we have $a(1) = 1$ and $b(1) = c(1) = 0$, we get

$$x \cdot \frac{d\mu}{dx} = A \cdot \mu + B - C \cdot \mu^2,$$

where A and B are the same as in the previous proof and $C \stackrel{\text{def}}{=} c'(1)$.

For $x \to \infty$, we have $\mu(x) \to 0$, so $\mu'(x) \to 0$, and thus $B = 0$ and

$$x \cdot \frac{d\mu}{dx} = A \cdot \mu - C \cdot \mu^2,$$

i.e.,

$$\frac{d\mu}{B \cdot \mu - C \cdot \mu^2} = \frac{dx}{x}. \tag{5}$$

As we have shown in the proof of Proposition 1, we cannot have $C = 0$, so $C \neq 0$. One can easily see that

$$\frac{1}{\mu - \dfrac{B}{C}} - \frac{1}{\mu} = \frac{\dfrac{B}{C}}{\mu \cdot \left(\mu - \dfrac{B}{C}\right)} = \frac{-B}{B \cdot \mu - C \cdot \mu^2}.$$

Thus, by multiplying both sides of equality (5) by $-B$, we get

$$\frac{d\mu}{\mu - \dfrac{B}{C}} - \frac{d\mu}{\mu} = -B \cdot \frac{dx}{x}.$$

Integrating both sides, we get

$$\ln\left(\mu(x) - \frac{B}{C}\right) - \ln(\mu) = -B \cdot \ln(x) + c_0.$$

By applying $\exp(z)$ to both sides, we get

$$\frac{\mu(x) - \dfrac{B}{C}}{\mu(x)} = C_0 \cdot x^{-B}$$

8

for some constant C_0, i.e.,

$$1 - \frac{B/C}{\mu} = C_0 \cdot x^{-B},$$

hence

$$\frac{B/C}{\mu} = 1 - C_0 \cdot x^{-B}$$

and

$$\mu(x) = \frac{B/C}{1 - C_0 \cdot x^{-B}}.$$

From the condition that $\mu(0) = 1$, we conclude that $B < 0$ and $B/C = 1$. From the condition that $\mu(x) \leq 1$, we conclude that $C_0 < 0$. Thus, we get the desired formula

$$\mu(x) = \frac{1}{1 + |C_0| \cdot x^{|B|}}.$$

The proposition is proven.

5.3 Proof of Proposition 3

For constant functions the statement is trivial, since every constant function is fractional-linear. Therefore, it is sufficient to prove for non-constant functions $h(x)$.

Similarly to the proof of Proposition 2, we can prove that the function $h(x)$ is differentiable. Let $x \in D$, and let λ and x_0 from an open neighborhood of 1 and 0 respectively be such that $\lambda \cdot x \in D$ and $x + x_0 \in D$. Since the function $h(x)$ is h-scale-invariant, there exist fractional-linear transformations for which

$$h(x + x_0) = \frac{a(x_0) \cdot h(x) + b(x_0)}{1 + c(x_0) \cdot h(x)} \tag{6}$$

and

$$h(\lambda \cdot x) = \frac{d(\lambda) \cdot h(x) + e(\lambda)}{1 + f(\lambda) \cdot h(x)}. \tag{7}$$

Similarly to the proof of Proposition 2, we can prove that the functions $a(x_0)$, ..., are differentiable. Similar to the proof of Proposition 2, we can differentiate the formula (7) with respect to λ and take $\lambda = 1$, then we get:

$$x \cdot h' = D \cdot h + E - F \cdot h^2. \tag{8}$$

Similarly, differentiating the formula (6) with respect to x_0 and taking $x_0 = 0$, we get:

$$h' = A \cdot h + B - C \cdot h^2. \tag{9}$$

Let us consider two cases: $C \neq 0$ and $C = 0$.

Let us first consider the case when $C \neq 0$. By completing the square, we get $h' = A \cdot h + B - C \cdot h^2 = \widehat{A} - C \cdot (h - h_0)^2$ for some $\widehat{A}$ and h_0, i.e.,

$$h' = \widehat{A} - C \cdot H^2, \tag{10}$$

where $H \stackrel{\text{def}}{=} h - h_0$. Substituting $h = H + h_0$ into the right-hand of the formula (8), we conclude that

$$x \cdot h' = \widehat{D} \cdot H + \widehat{E} - F \cdot H^2 \tag{11}$$

for some constants $\widehat{D}$ and $\widehat{E}$. Dividing (11) by (10), we get

$$x = \frac{\widehat{D} \cdot H + \widehat{E} - F \cdot H^2}{\widehat{A} - C \cdot H^2}, \tag{12}$$

so

$$\frac{dx}{dH} = \frac{(\widehat{D} - 2F \cdot H) \cdot (\widehat{A} - C \cdot H^2) - (\widehat{D} \cdot H + \widehat{E} - F \cdot H^2) \cdot (-2C \cdot H)}{(\widehat{A} - C \cdot H^2)^2} =$$

$$\frac{\widehat{A} \cdot \widehat{D} - 2(\widehat{A} \cdot F - C \cdot \widehat{E}) \cdot H + C \cdot \widehat{D} \cdot H^2}{(\widehat{A} - C \cdot H^2)^2}. \tag{13}$$

On the other hand,

$$\frac{dx}{dH} = \frac{1}{\dfrac{dH}{dx}} = \frac{1}{\widehat{A} - C \cdot H^2}. \tag{14}$$

The right-hand sides of the formulas (13) and (14) must be equal, so for all H, we have

$$\widehat{A} \cdot \widehat{D} - 2(\widehat{A} \cdot F - C \cdot \widehat{E}) \cdot H + C \cdot \widehat{D} \cdot H^2 = \widehat{A} - C \cdot H^2.$$

Since the two polynomials of H are equal, the coefficients at 1, H, and H^2 must coincide.

Comparing the coefficients at H^2, we get $C \cdot \widehat{D} = -C$. Since $C \neq 0$, we conclude that $\widehat{D} = -1$. Comparing the coefficients at 1, we get $\widehat{A} \cdot \widehat{D} = \widehat{A}$, i.e., $-\widehat{A} = \widehat{A}$ and thus $\widehat{A} = 0$. Comparing the coefficients at H and taking into account that $\widehat{A} = 0$, we get $0 = \widehat{A} \cdot F - C \cdot \widehat{E} = -C \cdot \widehat{E}$. Since $C \neq 0$, this implies $\widehat{E} = 0$. So, the formula (12) takes the form

$$x = \frac{\widehat{D} \cdot H - F \cdot H^2}{-C \cdot H^2} = \frac{\widehat{D} - F \cdot H}{-C \cdot H}.$$

Thus x is a fractional linear function of H, hence H (and therefore $h = H + h_0$) is also a fractional linear function of x.

Let us now consider the case when $C = 0$. In this case, $h' = A \cdot h + B$ and $x \cdot h' = D \cdot h + E - F \cdot h^2$, thus

$$x = \frac{x \cdot h'}{h'} = \frac{D \cdot h + E - F \cdot h^2}{A \cdot h + B}.$$

If $F = 0$, then x is a fractional linear function of $h(x)$ and hence, h is also a fractional-linear function of x.

10

So, it is sufficient to consider the case when $F \neq 0$. In this case, by completing the square, we can find constants $\widehat{D}$, h_0, and $\widehat{B}$ for which, for $H = h - h_0$, we have

$$x \cdot h' = D \cdot h + E - F \cdot h^2 = \widetilde{D} - F \cdot H^2 \tag{15}$$

and

$$h' = A \cdot h + B = A \cdot H + \widehat{B}. \tag{16}$$

Dividing (15) by (16), we have

$$x = \frac{\widetilde{D} - F \cdot H^2}{A \cdot H + \widehat{B}}. \tag{17}$$

Thus,

$$\frac{dx}{dH} = \frac{(-2F \cdot H) \cdot (A \cdot H + \widehat{B}) - (\widehat{D} - F \cdot H^2) \cdot A}{(A \cdot H + \widehat{B})^2}$$

$$= \frac{-A \cdot \widehat{D} - 2\widehat{B} \cdot F \cdot H - A \cdot F \cdot H^2}{(A \cdot H + \widehat{B})^2}.$$

On the other hand,

$$\frac{dx}{dH} = \frac{1}{\dfrac{dH}{dx}} = \frac{1}{A \cdot H + \widehat{B}}.$$

By equating the two expressions for the derivative and multiplying both sides by $(A \cdot H + \widehat{B})^2$, we conclude that

$$-A \cdot \widehat{D} - 2\widehat{B} \cdot F \cdot H - A \cdot F \cdot H^2 = A \cdot H + \widehat{B},$$

thus $A \cdot F = 0$, $A = -2\widehat{B} \cdot F$, and $-A \cdot \widehat{D} = \widehat{B}$. If $A = 0$, then we have $\widehat{B} = 0$, so $h' = 0$ and h is a constant – but we consider the case when the function $h(x)$ is not a constant. Thus, $A \neq 0$, hence $F = 0$, and the formula (17) describes x as a fractional-linear function of H.

In both cases $C \neq 0$ and $C = 0$, we obtain an expression of x in terms of H (hence h) that is fractional linear. Since the inverse of a fractional linear is fractional linear, the function $h(x)$ is also fractional linear.

The proposition is proven.

Acknowledgments

References

[1] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[2] O. Csiszar and J. Dombi, "Generator-based modifiers and membership functions in nilpotent operator systems", *Proceedings of the IEEE International Work Conference on Bioinspired Intelligence IWOBI'2019*, Budapest, Hungary, July 3–5, 2019, pp. 99–105.

[3] J. Dombi and O. Csiszar, "Implications in bounded systems", *Information Sciences*, 2014, Vol. 283, pp. 229–240.

[4] J. Dombi and O. Csiszar, "The general nilpotent operator system", *Fuzzy Sets and Systems*, 2015, Vol. 261, pp. 1–19.

[5] J. Dombi and O. Csiszar, "Equivalence operators in nilpotent systems", *Fuzzy Sets and Systems*, 2016, Vol. 299, pp. 113–129.

[6] J. Dombi and T. Szépe, "Arithmetic-based fuzzy control", *Iranian Journal of Fuzzy Systems*, 2017, Vol. 14, No. 4, pp. 51–66.

[7] J. Dombi and A. Hussain, "Interval type-2 fuzzy control using the distending function", In: A. J. Tallón-Ballesteros (ed.), *Proceedings of the 5th International Conference on Fuzzy Systems*, Kitakyushu, Japan, October 18–21, 2019, pp. 705–714.

[8] J. Dombi and A. Hussain, "Data-driven arithmetic fuzzy control using the distending function", *Proceedings of the International Conference on Human Interaction and Emerging Technologies IHIET'2019*, Nice, France, August 22–24, 2019, pp. 215–221.

[9] J. Dombi and A. Hussain, "A new approach to fuzzy control using the distending function", *Journal of Process Control*, 2020, Vol. 86, pp. 16–29.

[10] V. M. Guillemin and S. Sternberg, "An algebraic model of transitive differential geometry", *Bulletin of the American Matjematical Society*, 1964, Vol. 70, No. 1, pp. 16–47.

[11] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[12] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[13] H. T. Nguyen and V. Kreinovich, *Applications of Continuous Mathematics to Computer Science*, Kluwer, Dordrecht, 1997.

[14] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[15] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[16] I. M. Singer and S. Sternberg, Infinite groupsof Lie and Cartan, Part I, *Journal d'Analyse Mathematique*, 1965, Vol. 15, pp. 1–113.

[17] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.