Thai **J**ournal of **M**athematics : (2019) 41–55
Special Issue: Structural Change Modeling
and Optimization in Econometrics 2018
http://thaijmath.in.cmu.ac.th
ISSN 1686-0209

Translating Discrete Estimates into a Less Detailed Scale: An Optimal Approach

Thongchai Dumrongpokaphan[†], **Olga Kosheleva**[‡] **and Vladik Kreinovich**[‡,1]

[†]Faculty of Economics, Chiang Mai University
Chiang Mai 50200, Thailand
e-mail : `tcd43@hotmail.com`

[‡]University of Texas at El Paso, USA
e-mail : `olgak@utep.edu (O. Kosheleva)`
`vladik@utep.edu (V. Kreinovich)`

Abstract : In many practical situations, we use estimates that experts make on a 0-to-n scale. For example, to estimate the quality of a lecturer, we ask each student to evaluate this quality by selecting an integer from 0 to n. Each such estimate may be subjective; so, to increase the estimates' reliability, it is desirable to combine several estimates of the corresponding quality. Sometimes, different estimators use slightly different scales: e.g., one estimator uses a scale from 0 to $n+1$, and another estimator uses a scale from 0 to n. In such situations, it is desirable to translate these estimates to the same scale, i.e., to translate the first estimator's estimates into the 0-to-n scale. There are many possible translations of this type. In this paper, we find a translation which is optimal under a reasonable optimality criterion.

Keywords : discrete scale; Likert scale; translation between discrete scales; optimality; symmetry.
2010 Mathematics Subject Classification : 91E45; 97N70; 39A12; 68R99.

---

[0]This work was supported in part by the Chiang Mai University, by the National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence), and by the University of Ostrava.
[1]`Corresponding author email.`

1 Formulation of the Problem

Need for discrete estimates. Discrete estimates are important in many real-life problems. These estimates are important in evaluating situations and in making decisions.

When we evaluate a given situation, in addition to numerical characteristics, we often use discrete estimates; for example:

- to estimate the quality of a hotel stay, in addition to price and wait times at reception, we also ask customers to estimate, on a certain discrete scale, how much they liked the hotel in general and/or different aspects of the hotel service;
- students evaluate the quality of their instructors by using a discrete scale – e.g., from 0 to 5, etc.

In decision making, discrete estimates are also important. Indeed, a large number of decisions are made by experts; for example:

- a medical doctor prescribes a certain treatment,
- a skilled driver decides how much to break if a road situation changes,
- an investor decides whether to re-balance her investment.

In all these situations, the expert bases his/her decisions not only on numerical values of the corresponding quantities, but also on a discrete estimate.

- When encountering a skin inflammation, the doctor takes into account whether the area of this inflammation is small, medium, or large.
- A driver makes different decisions depending on whether the car in front slowed down a little bit, some, or drastically.
- An investor bases his/her decision on whether the prices of different stocks increased a lot, increases somewhat, increased slightly, or decreased (and decreased to what extent).

In all these cases, we use a discrete (Likert-type) scale. In some cases, each element of the corresponding scale has a natural-language description – such as "small", "medium", "large", etc. In other cases, we simply ask the users to estimate their opinion on a scale, e.g., from 0 to 10, but we only assign natural-language explanations to the boundary values (0 and 10), and not to intermediate degrees. In all these cases, we have a scale whose elements can be numbered in increasing order: 0, 1, ..., up to a certain natural number n.

Comment. For situations when each element of a scale has a natural-language description, there is a special methodology for dealing with such cases – the technique of fuzzy logic; see, e.g., [1, 2, 3, 4, 5, 6].

Need to translate discrete estimates into a different scale. The opinion of each individual expert may be subjective and biased. To get a more accurate understanding of the current situation and/or a more reliable decision, we must combine the opinion of several different experts.

One of the difficulties in such combination is that different experts use, in general, different scales. This is a known phenomenon in psychology: there is a so-called 7 ± 2 phenomenon (see, e.g., [7, 8]), according to which each person is most comfortable with a certain size of scale – from scale with $7 - 2 = 5$ elements to a scale with $7 + 2 = 9$ elements.

Thus, to combine different expert estimates, it is necessary to transform them into a single scale.

We need to translate to a less detailed scale. The fewer the number of elements in a scale, the less information is contained in the corresponding estimate. For example, if we have a binary "yes"-"no" scale, with two elements, the corresponding estimate clearly contains less information that the more detailed scale, e.g., a scale "absolutely yes" – "somewhat confident that yes" – "neutral" – "somewhat confident that no" – "no". From this viewpoint, to translate into a more detailed scale is simpler: we will simply leave out some intermediate values. A more challenging problem is how to translate an estimate into a less detailed scale.

It is important to transform discrete estimates to a scale that has one fewer element. In general, we have many experts, with different numbers of elements in their corresponding scales. Intuitively, the larger the difference between numbers of elements in two scales, the more difficult it is to translate between these two scales. It is therefore reasonable to concentrate on how to translate from a scale to a closest scale – namely, the scale that has one fewer element.

Once we learn how to do it, we can handle larger differences as well: namely, we can first translate all 9-element-scale estimates into an 8-element scale. Then, we can translate these newly formed 8-element-scale estimates and the original 8-element-scale estimates into a 7-element scale, etc.

What we do in this paper. In this paper, after overviewing different technique for such a translation, we formulate the problem of selecting a translation as an optimization problem. Then, we solve this optimization problem – and thus, describe which translations are optimal.

2 How to Translate: A Currently Used Straightforward Approach

Numerical methods: main idea. Computers have been designed to deal with numbers, not with elements of a scale. Thus, to process scale-based expert information, a natural idea is to translate such information into numbers. For example, we can use numbers from the interval $[0, 1]$.

Once such a translation is selected, there is a simply way to solve our translation problem: to see what i on a scale from 0 to $n + 1$ means in a 0-to-n scale, we simply take the number corresponding to i on a 0-to-$(n+1)$ scale and find the closest of the numbers describing the 0-to-n scale.

To utilize the numerical approach, all we need to do is to decide, for each element i on each 0-to-n scale, which number $v_{i,n} \in [0,1]$ to assign to this element. The only requirement is that smaller elements should be described by smaller numbers.

Definition 1. *By a* numerical equivalent*, we mean a sequence of numbers $v_{i,n}$ define for all $n \geq 1$ and for all i from 0 to n such that for every n and for every $i < n$, we have $v_{i,n} < v_{i+1,n}$.*

Definition 2. *Let a numerical equivalent $v_{i,n}$ be given. Then, for each n and for each $i \leq n+1$, by the corresponding n-*translation $t_n(i)$*, we mean an integer j for which the value $v_{j,n}$ is the closest to $v_{i,n+1}$:*

$$|v_{t_n(i),n} - v_{i,n+1}| = \min_{0 \leq j \leq n} |v_{j,n} - v_{i,n+1}|.$$

Straightforward approach. In the straightforward approach, we associate the smallest element of the scale with 0, the largest with 1, and we made all the other values equally spaced, i.e., we take $v_{i,n} = \dfrac{i}{n}$.

Definition 3. *By a* straightforward approach*, we mean a numerical equivalent*

$$v_{i,n} = \frac{i}{n}.$$

Comment. This is the most widely used numerical translation of a scale. This is, e.g., how scales are translated into numbers in fuzzy logic [1, 2, 3, 4, 5, 6].

Resulting translation into a less detailed scale. Now that we know the values $v_{i,n}$ corresponding to the straightforward approach, we can determine the corresponding translation into a less detailed scale. The results are somewhat different for even and odd n.

Proposition 1. *For even n, for the straightforward approach, the corresponding n-translation has the following form:*

- *for $i < (n+1)/2$, we have $t_n(i) = i$, and*
- *for $i > (n+1)/2$, we have $t_n(i) = i - 1$.*

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

Comment. In other words, for $n = 2k$:

- values 0, ..., $k-1$ on the 0-to-$(2k+1)$ scale translate into themselves,
- values k and $k+1$ are both translated into k, and

- values $k+2$, $k+3$, …, $n+1$ are translated, accordingly, into $k+1$, $k+2$, …, n.

Proposition 2. *For odd n, for the straightforward approach, the corresponding n-translation has the following form:*

- *for $i < (n+1)/2$, we have $t_n(i) = i$,*
- *for $i = (n+1)/2$, we have either $t_n(i) = i$ or $t_n(i) = i - 1$, and*
- *for $i > (n+1)/2$, we have $t_n(i) = i - 1$.*

Comment. In other words, for $n = 2k - 1$:

- values 0, ..., $k-1$ on the 0-to-$2k$ scale translate into themselves,
- value k can be translated either into k or into $k-1$, and
- values $k+1$, $k+2$, …, $n+1$ are translated, accordingly, into k, $k+1$, …, n.

Comment. When applying several consequent transitions from a scale to a less detailed one is that the result of the consequent n-translations may differ from what we would get if we translate directly. Let us give an example. Suppose that we start with the 0-to-5 scale in which, in the straightforward approach, the numerical equivalents are 0, 1/5, 2/5, 3/5, 4/5, and 1. We want to translate it into the 0-to-2 scale, in which the numerical equivalents are 0, 1/2, and 1.

There are two ways to perform this n-translation:

- first, we can first go from the 0-to-5 scale to the 0-to-4 scale, then to the 0-to-3-scale, and finally, to the 0-to-2 scale;
- alternatively, we can directly go from the 0-to-5 scale to the 0-to-2 scale – by assigning to each numerical value from the 0-to-5 scale the closest numerical value on the 0-to-2 scale.

Let us trace how the grade 1 on the 0-to-5 scale – whose numerical equivalent is 1/5 – gets translated in both approaches.

- In the first approach, according to our results (as expressed by Propositions 1 and 2), 1/5 gets translated into 1/4, then into 1/3, and finally, into 1/2.
- However, in the second approach, since 1/5 is closer to 0 than to 1/2 or to 1, the value 1/5 will be translated into 0 – and not to 1/2 as in the first approach.

3 Maximum Entropy Approach

Main idea. For each n, the only restriction that we have on the corresponding values $v_{0,n}, v_{1,n}, \ldots, v_{n,n}$ is that $v_{0,n} < v_{1,n} < \ldots < v_{n,n}$. Let V_n denote the set of all the tuples $v_n = (v_{0,n}, v_{1,n}, \ldots, v_{n,n})$ that satisfy this property.

If we knew the probability distribution $\rho(v_n)$ on this set V_n, then it makes sense to select a tuple $\overline{v}_n \in V_n$ for which the mean square deviation is the smallest possible, i.e., the tuple $\overline{v}_n$ that minimizes the integral

$$\int \rho(v_n) \cdot (v_n - \overline{v}_n)^2 \, dv_n.$$

In general, in statistics, it is well known that this minimum is attained at the mean value $\overline{v}_n = \int \rho(v_n) \cdot v_n \, dv_n$; see, e.g., [9].

To follow this approach, we need to select a distribution on the corresponding set of tuples. Since we have no reason to believe that some tuples are more probable than others, it makes sense to assume that all the tuples from V_n are equally probable, i.e., that we have a uniform distribution on the set of all such tuples. This natural idea is known as *Laplace Indeterminacy Principle*, and it is a particular case of a general Maximum Entropy approach; see, e.g., [10].

Known auxiliary result. The mean values of $v_{i,n}$ with respect to the uniform distribution on the set V_n of all the tuples are known (see, e.g., [11, 12, 13, 14, 15, 16, 17, 18, 19]): they are $v_{i,n} = \dfrac{i+1}{n+2}$. So, we arrive at the following definition.

Definition 4. *By a* maximum entropy approach*, we mean a numerical equivalent*

$$v_{i,n} = \frac{i+1}{n+2}.$$

Comment. It is worth mentioning that exactly the same tuples appear if we select the tuple v_n which is the most robust (in some reasonable sense); see, e.g., [15].

Resulting n-translation. It turns out that for the maximum entropy approach, we get the exact same n-translation as for the straightforward approach:

Proposition 3. *For even n, for the maximum entropy approach, the corresponding n-translation has the following form:*

- *for $i < (n+1)/2$, we have $t_n(i) = i$, and*
- *for $i > (n+1)/2$, we have $t_n(i) = i - 1$.*

Proposition 4. *For odd n, for the maximum, entropy approach, the corresponding n-translation has the following form:*

- *for $i < (n+1)/2$, we have $t_n(i) = i$,*
- *for $i = (n+1)/2$, we have either $t_n(i) = i$ or $t_n(i) = i - 1$, and*
- *for $i > (n+1)/2$, we have $t_n(i) = i - 1$.*

4 General Approach

Why we need a general approach. All we have is values on a scale. So it is more natural to deal directly with numbers on a scale and not artificially add numbers to the clearly non-numeric data. This is what we will do in this section.

Motivations and the resulting definitions. For every n, we need to describe a mapping $t_n : \{0, 1, \ldots, n+1\} \to \{0, 1, \ldots, n\}$. It is reasonable to require:

- that the worst case gets translated into the worst case, i.e., that $t_n(0) = 0$, and
- that the best case gets translated into the best case, i.e., that $t_n(n+1) = n$.

Definition 5. *For every $n \geq 1$, by an* n-translation*, we mean a mapping*

$$t_n : \{0, 1, \ldots, n+1\} \to \{0, 1, \ldots, n\}$$

for which $t_n(0) = 0$ and $t_n(n+1) = n$.

It is also reasonable to require that small changes in the input to the n-translation function cause small changes in the output. Since each change can be described as a superposition of changes by 1, it is sufficient to formulate this "continuity" property for the case when the change means adding or subtracting 1.

Definition 6. *We say that the values i and i' are* close *if $|i - i'| \leq 1$.*

Definition 7. *We say that a n-translation t_n is* continuous *if whenever i and i' are close, the values $t_n(i)$ and $t_n(i')$ are also close.*

We want to select one of the possible continuous n-translations. There are several different continuous n-translations. For example, for $n = 2$, in addition to $t_2(0) = 0$ and $t_2(3) = 2$, we can have at least two different options:

- we can have $t_2(1) = 0$ and $t_2(2) = 1$, or
- we can have $t_2(1) = 1$ and $t_2(2) = 1$, or
- we can have $t_2(1) = 1$ and $t_2(2) = 2$.

In all these cases, we have a continuous n-translation function.

We therefore need to select one of the possible continuous n-translations.

We want an optimal n-translation – but how do we describe this in precise terms? Of course, we want to select the "best" (optimal) one – the best in terms of the resulting applications. To perform such selection, we need to describe what "the best" means.

In different practical situations, however, optimal may mean different things. So, ideally, instead of selecting a single optimality criterion, we should consider all possible reasonable optimality criteria.

What is an optimality criterion? Usually, we select some objective function $J(a)$, and consider an alternative a to be better than alternative a' if $J(a) > J(a')$ (or, alternatively, $J(a) < J(a')$). However, this is not the most general description of optimality. For example, if we select an algorithm a with the best possible worst-case complexity $J(a)$, and there are several such algorithms, then we can use this non-uniqueness to optimize something else, e.g., the average-case complexity $A(a)$. In this case, the actual optimality criterion that we use to select an algorithm is no longer a numerical one, it is more complicated. Namely, a is better than a' if:

- either $J(a) < J(a')$,
- or $J(a) = J(a')$ and $A(a) < A(a')$.

If we still have several algorithms which are equally good with respect to this complex criterion, we can use this non-uniqueness to optimize something else, etc., until we get to the point when we are left with exactly one optimal alternative.

In general, what we want from an optimality criterion is that it allows us, for every two alternatives a and a',

- either to select one of them,
- or to conclude that a and a' are equivalent,
- or maybe to conclude that a and a' are incompatible.

Let us denote the case when a is better than a' or of the sane quality by $a \preceq a'$.

Clearly, each alternative a is of the same quality as itself, so we must have $a \preceq a$. In mathematical terms, this means that the relation $\preceq$ is *reflexive*.

If a is better than a' and a' is better than a'', this means that a is better than a''. In other words, if $a \preceq a'$ and $a' \preceq a''$, then we should have $a \preceq a''$. In mathematical terms, this means that the relation $\preceq$ is *transitive*.

Reflexive transitive relations are known as *pre-orders*. They are not necessarily order: an order has an additional property that if $a \preceq a'$ and $a' \preceq a$, then $a = a'$. For optimality criterion, this is not necessarily true: nothing wrong with having two different alternatives which are equally bad. Thus, we arrive at the following definition; see, e.g., [20].

Definition 8. *By an* optimality criterion *on a set A, we mean a pre-order relation $\preceq$ on this set, i.e., a relation which satisfies the following two properties:*

- *reflexivity $a \preceq a$, and*
- *transitivity: if $a \preceq a'$ and $a' \preceq a''$, then $a \preceq a''$.*

Definition 9. *Let $\preceq$ be an optimality criterion on a set A. We say that the alternative $a_{\rm opt}$ is* optimal *if $a_{\rm opt} \preceq a$ for all $a \in A$.*

When is an optimality criterion final? As we argued earlier, if an optimality criterion selects several different alternatives as optimal, this means that this criterion is not final: we can use this non-uniqueness to optimize something else and thus, we modify the original criterion. Such a modification continues until we get an optimality criterion that has exactly one optimal alternative. Thus, we arrive at the following definition.

Definition 10. *We say that an optimality criterion $\preceq$ is* final *if there exists exactly one alternative which is optimal with respect to this criterion.*

Symmetries: main idea. In many practical situations, there exist some natural symmetries. In such situations, it is reasonable to require that the optimality criterion be invariant under this transformation: if $a \preceq a'$ and T is the corresponding symmetry, then we should have $T(a) \preceq T(a')$.

Symmetries: our case. For each scale, we can select two different directions; for example:

- we can mark the awful professor by 0 and the best professor by 5; in this case, the number describes the quality of a professor;
- alternatively, we can mark the awful professor by 5 and the best professor by 0; in this case, the number describes the size of this professor's limitations.

In general, such a scale reversal, from the original values 0, 1, ..., $n-1$, n to the new values n, $n-1$, ..., 1, 0 can be described by a formula $i \to R_n(i) \stackrel{\text{def}}{=} n - i$.

If in the original scale, we have a n-translation $t_n : \{0, 1, \ldots, n+1\} \to \{0, 1, \ldots, n\}$, then in the reverse scales, this n-translation takes the following form:

- first, we reverse the given value i on the new scale into a reversed value
$$R_{n+1}(i) = (n+1) - i;$$
- then, we apply the original n-translation, and get the value $t_n(R_{n+1}(i))$ in the old scale,
- finally, we reverse this value, to get to the new scale; this results in the value
$$R_n(t_n(R_{n+1}(i))).$$

Thus, we arrive at the following definition.

Definition 10. *For each n, by $R_n(i)$, we mean the value $n - i$. For each n-translation t_n, by its reversal, we mean a n-translation $R(t_n)$ which is defined as follows: $(R(t_n))(i) = R_n(t_n(R_{n+1}(i)))$.*

Definition 11. *Let n be given. We say that the optimality criterion on the set of all possible continuous n-translations is* invariant *if $a \preceq a'$ always implies*
$$R(a) \preceq R(a').$$

First result: case of even n. Invariance property enables us to find the optimal n-translation for the case when the value n is even:

Proposition 5. *For even n, for every invariant final optimality criterion on the set of all continuous n-translations, the optimal n-translation is the one corresponding to the straightforward approach.*

Case of off n: discussion and the main result. For the case when n is odd, the above approach does not work:

Proposition 6. *Let n be an odd number. Then, no final optimality criterion on the set of all continuous n-translations is invariant.*

Discussion. This result shows that for odd n, it is not possible to have a unique optimal n-translation. Thus, there must be several optimal n-translations. Ideally, the smaller the number of optimal n-translations, the better. Since we cannot have a single optimal n-translation, let us thus consider the possibility to have two different optimal n-translations.

We cannot have these two optimal n-translations to be equal to each other, but at least we should require that they be as close to each other as possible. From the mathematical viewpoint, n-translations are functions. The two functions are equal if for each input x, their values are equal. We cannot have *all* the values equal, so the next best thing is to have these values equal for all but one inputs x. Thus, we arrive at the following definitions.

Definition 12. *We say that two n-translations t_n and s_n are* almost equal *if their values coincide for all but one value i.*

Definition 13. *We say that the optimality criterion on the class of all continuous n-translations is* almost final *if for this criterion, there exist exactly two optimal n-translations, and these two n-translations are almost equal.*

Now, we can determine the optimal n-translation.

Proposition 7. *For odd n, for every invariant almost final optimality criterion on the set of all continuous n-translations, the optimal n-translations are the ones corresponding to the straightforward approach.*

5 Proofs

Proof of Propositions 1 and 2. For each i, let us find the fraction $v_{j,n} = \dfrac{j}{n}$ which is the closest to the original value $v_{i,n+1} = \dfrac{i}{n+1}$.

For $i = 0$ and $i = n + 1$, this selection is clear, since in these cases, we can get not just the closest values, we can get exact equality: $v_{0,n+1} = v_{0,n} = 0$ and $v_{n+1,n+1} = v_{n,n} = 1$. So, it is sufficient to consider the case when $0 < i \le n$.

In this case, we always have $\frac{i}{n+1} < \frac{i}{n}$. Let us show that always have $\frac{i-1}{n} < \frac{i}{n+1}$. Indeed, if we multiply both sides by $n \cdot (n+1)$, we get an equivalent inequality $(i-1) \cdot (n+1) < i \cdot n$, i.e., $i \cdot n + i - n - 1 < i \cdot n$, which is equivalent to the true inequality $i < n+1$.

Since $\frac{i-1}{n} < \frac{i}{n+1} < \frac{i}{n}$, the closest fraction to $\frac{i}{n+1}$ is either $\frac{i-1}{n}$, or $\frac{i}{n}$. To find out which is closer, we need to compare the corresponding distances: the value $j = i$ is closer if and only if

$$\frac{i}{n} - \frac{i}{n+1} \leq \frac{i}{n+1} - \frac{i-1}{n}.$$

Multiplying both sides by $n \cdot (n+1)$, we get an equivalent inequality

$$i \cdot (n+1) - i \cdot n \leq i \cdot n - (i-1) \cdot (n+1).$$

If we open parentheses and cancel the terms $i \cdot n$ and $-i \cdot n$ in both sides, we get an equivalent inequality $i \leq -i + n + 1$. Moving i to the right-hand side, we get an equivalent inequality $2i \leq n+1$, i.e., $i \leq \frac{n+1}{2}$. Thus:

- when $i < \frac{n+1}{2}$, the value $v_{i,n}$ is closer,
- when $i = \frac{n+1}{2}$, both values $v_{i,n}$ and $v_{i-1,n}$ are equally close, and
- when $i > \frac{n+1}{2}$, the value $v_{i-1,n}$ is closer.

Thus, the propositions are proven.

Proof of Proposition 3 and 4 is similar to the proof of Propositions 1 and 2.

Proof of Propositions 5 and 6.

1°. By definition of an n-translation, we have $t_n(0) = 0$ and $t_n(n+1) = n$. The difference $t_n(n+1) - t_n(0)$ can be represented as the sum $n+1$ differences between neighboring values $t_n(i+1)$ and $t_n(i)$:

$$n = t_n(n+1) - t_n(0) =$$

$$(t_n(1) - t_n(0)) + (t_n(2) - t_n(1)) + \ldots + (t_n(n+1) - t_n(n)).$$

Since we consider continuous n-translations, for each i, the neighboring difference $t_n(i+1) - t_n(i)$ is equal either to -1, or to 0, or to 1. In all three cases, this difference does not exceed 1. Thus, the value $n = t_n(n+1) - t_n(0)$ is smaller than or equal to the number of transitions from $t_n(i)$ to $t_n(i+1)$ for which $t_n(i+1) - t_n(i) = 1$. Hence, we must have at least n transitions for which $t_n(i)$ increases by 1 as we go from the previous value i to the next value $i+1$.

These n transitions already provide the sum n, so the only way to keep the overall sum equal to n is when the single remaining neighboring difference is equal to 0.

Let j denote the index of this remaining 0-valued difference, i.e., the value j for which $t_n(j+1) - t_n(j) = 0$. In this case, the n-translation has the following form.

- For $i \le j$, we have
$$t_n(i) = t_n(i) - t_n(0) =$$
$$(t_n(1) - t_n(0)) + (t_n(2) - t_n(1)) + \ldots + (t_n(i) - t_n(i-1)).$$
All the terms in the right-hand side are 1s, and there are i such terms, so we conclude that
$$t_n(i) = i.$$

- For $i > j$, we have
$$t_n(i) - t_n(j) =$$
$$(t_n(j+1) - t_n(j)) + (t_n(j+2) - t_n(j+1)) + \ldots + (t_n(i) - t_n(i-1)).$$
In this case, the first term in the right-hand side is 0, and all others – and there are $i - j - 1$ of them – are equal to 1. Thus, we have $t_n(i) - t_n(j) = i - j - 1$ and hence,
$$t_n(i) = t_n(j) + (i - j - 1) = j + (i - j - 1) = i - 1.$$

2°. Let us now prove that the optimal n-translation a should itself be reverse-invariant, i.e., we should have $R(a) = a$.

Indeed, by definition, optimality means that $a \preceq t$ for any continuous n-translation t. In particular, this n-translation t can have the form $t = R(s)$ for some other n-translation s, so we can conclude that $a \preceq R(s)$ for all continuous n-translations s. Since the optimality criterion is invariant, we conclude that $R(a) \preceq R(R(s))$.

One can easily check that the reversal, when applied twice, returns back to the original n-translation, i.e., that we have $R(R(s)) = s$ for all s. Thus, the above property $R(a) \preceq R(R(s))$ implies that $R(a) \preceq s$ for all continuous n-translations s. This means that the reversed n-translation $R(a)$ is also optimal.

However, our optimality criterion is final – which means that there is only one optimal n-translation. Thus, we must have $R(a) = a$.

3°. We know that in each continuous n-translation, there is a single pair of neighboring values $(j, j+1)$ for which $t_n(j) = t_n(j+1)$. In particular, this is true for the optimal n-translation a: there is a single pair of neighboring values $(j, j+1)$ for which $a(j) = a(j+1)$.

In the reversed n-translation $R(a)$, the reversed values $R_{n+1}(j) = n + 1 - j$ and $R_{n+1}(j+1) = n+1-(j+1) = n-j$ have the same property: $(R(a))(n-j) =$

$(R(a))(n-j+1)$. Since $R(a) = a$, the corresponding pair $(n-j, n-j+1)$ must coincide with the original pair $(j, j+1)$. Thus, we must have $n-j = j$, i.e., $n = 2j$.

When n is even, this means that we must have $j = n/2$ – in which case the n-translation described in Part 1 of this proof coincides with the n-translation corresponding to the straightforward approach. This proves Proposition 5.

When n is odd, the equality $n = 2j$, where j is an integer index, is not possible. This proves Proposition 6.

Proof of Proposition 7. In the proof of Propositions 5 and 6, we have described the general form of a continuous n-translation. Each such n-translation is uniquely described by a pair of neighboring indices $(j, j+1)$ for which $t_n(j) = t_n(j+1)$.

One can easily see that the two n-translations corresponding to pairs $(j, j+1)$ and $(j', j'+1)$ are almost equal if the corresponding value j and j' differ by 1.

Similarly to the proof of Propositions 5 and 6, we can conclude that if an n-translation is optimal, then the reversed n-translation is also optimal. If for the original optimal n-translation, the pair of indices for which the n-translation leads to the same result is $(j, j+1)$, then for the reverse n-translation this pairs consists of the reversals $R_{n+1}(j) = n+1-j$ and $R_{n+1}(j+1) = n+1-(j+1) = n-j$. Thus, for the reverse n-translation, the corresponding pair is $(n-j, n-j+1)$.

Since the two optimal n-translations should be almost equal, the corresponding indices should differ by 1: $|(n-j)-j| = |n-2j| \leq 1$, or, equivalently, $-1 \leq 2j-n \leq 1$, i.e., $n-1 \leq 2j \leq n+1$. For odd n, i.e., for n of the type $n = 2k-1$, this implies $2k-2 \leq 2 \leq 2k$, i.e., $k-1 \leq j \leq k$. In both cases, we indeed have the n-translation corresponding to the straightforward approach. The proposition is proven.

Acknowledgement : The authors are very thankful to Michal Holčapek and Irina Perfilieva from the University of Ostrava for valuable discussions.

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

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

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

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

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

[7] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information", *Psychological Review*, 1956, Vol. 63, No. 2, pp. 81–97.

[8] S. K. Reed, *Cognition: Theories and application*, Wadsworth Cengage Learning, Belmont, California, 2010.

[9] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

[10] E. T. Jaynes and G. L. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.

[11] M. Ahsanullah, V. B. Nevzorov, and M. Shakil, *An Introduction to Order Statistics*, Atlantis Press, Paris, 2013.

[12] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*, Society of Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania, 2008.

[13] H. A. David and H. N. Nagaraja, *Order Statistics*, Wiley, New York, 2003.

[14] O. Kosheleva, V. Kreinovich, J. Lorkowski, and M. Osegueda, "How to transform partial order between degrees into numerical values", *Proceedings of the 2016 IEEE International Conferences on Systems, Man, and Cybernetics SMC'2016*, Budapest, Hungary, October 9–12, 2016.

[15] O. Kosheleva, V. Kreinovich, M. Osegueda Escobar, and K. Kato, "Towards the most robust way of assigning numerical degrees to ordered labels, with possible applications to dark matter and dark energy", *Proceedings of the 2016 Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2016*, El Paso, Texas, October 31 – November 4, 2016.

[16] J. Lorkowski and V. Kreinovich, "Interval and symmetry approaches to uncertainty – pioneered by Wiener – help explain seemingly irrational human behavior: a case study", *Proceedings of the 2014 Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2014*, Boston, Massachusetts, June 24–26, 2014.

[17] J. Lorkowski and V. Kreinovich, "Likert-type fuzzy uncertainty from a traditional decision making viewpoint: how symmetry helps explain human decision making (including seemingly irrational behavior)", *Applied and Computational Mathematics*, 2014, Vol. 13, No. 3, pp. 275–298.

[18] J. Lorkowski and V. Kreinovich, "Granularity helps explain seemingly irrational features of human decision making", In: W. Pedrycz and S.-M. Chen (eds.), *Granular Computing and Decision-Making: Interactive and Iterative Approaches*, Springer Verlag, Cham, Switzerland, 2015, pp. 1–31.

[19] J. Lorkowski and V. Kreinovich, "Fuzzy logic ideas can help in explaining Kahneman and Tversky's empirical decision weights", In: L. Zadeh et al. (Eds.), *Recent Developments and New Direction in Soft-Computing Foundations and Applications*, Springer Verlag, to appear.

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

(Received 18 November 2018)
(Accepted 23 January 2019)