Article

# Why Triangular Membership Functions Are So Efficient in F-Transform Applications: A Global Explanation to Supplement the Existing Local One

**Olga Kosheleva [1,‡], Vladik Kreinovich [1,†,‡]*, and Thach Ngoc Nguyen [2,‡]**

1 University of Texas at El Paso; olgak@utep.edu, vladik@utep.edu,
2 Banking University of Ho Chi Minh City, Vietnam; Thachnn@buh.edu.vn
* Correspondence: vladik@utep.edu; Tel.: +1-915-747-6951 (V.K.)
† Current address: Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
‡ These authors contributed equally to this work.

Version April 22, 2019 submitted to Preprints

**Abstract:** The main ideas of F-transform came from representing expert rules. It would be therefore re reasonable to expect that the more accurately the membership functions describe human reasoning, the more efficient will be the corresponding F-transform formulas. We know that an adequate description of our reasoning corresponds to complicated membership functions – however, somewhat surprisingly, most efficient applications of F-transform use the simplest possible triangular membership functions. There exist some explanations for this phenomenon which are based on local behavior of the signal. In this paper, we supplement this local explanation by a global one: namely, we prove that triangular membership functions are the only one that provide the accurate description of appropriate global characteristics of the signal.

**Keywords:** F-transform; triangular membership function; optimal global characteristics

---

## 1. Formulation of the Problem

**F-transforms: a brief reminder.** In many application areas, it turned to be very efficient to transform the original signal $x(t)$ into the values proportional to

$$x_i = \int A\left(\frac{t-t_i}{h}\right) \cdot x(t)\,dt,$$

where $t_i = t_0 + i \cdot h$ for appropriate $t_0$ and $h > 0$, and $A(t)$ is a non-negative function:

- which is equal to 0 outside the interval $[-1, 1]$,
- which, starting at $t = -1$, increases to 1 until it reaches $t = 0$,
- which then decreases to 0, and
- for which
$$\sum_i A\left(\frac{t-t_i}{h}\right) = 1$$
for all $t$; this last property is known as the *fuzzy partition property*.

This transformation is known as *F-transform*; see, e.g., [8,9,11–14].

This transform comes from the general fuzzy approach (see, e.g., [2,3,5,7,10,15]), namely, from the idea of describing imprecise (fuzzy) expert knowledge, of the type "if $t$ is close to $t_i$, then $x(t)$ is close to $x(t_i)$". From this viewpoint, the function $A(t)$ is a membership function that corresponds to the word "close".

**A somewhat unexpected empirical fact.** Intuitively, one would expect that the closer the function $A(t)$ to how we actually think, the more efficient would be the results. Empirical studies show that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2019 doi:10.20944/preprints201904.0253.v1



rather complex membership functions are needed to represent our reasoning; see, e.g., [10]. However, surprisingly, in many of these applications, very efficient results are obtained when we use a very simple triangular membership function $A(t) = 1 - |t|$. Why?

One possible "local" explanation – based on uncertainty – was proposed in [4]; however, not everyone was convinced, so this empirical fact still remains somewhat a mystery.

**What we do in this paper.** In this paper, we propose an alternative "global" explanation for this efficiency, an explanation based on the need to correctly reconstruct global characteristics of the signal.

## 2. Local Vs. Global Characteristics: Main Idea

**What we mean by local and global characteristics.** No measuring instrument can provide an instantaneous value of a physical quantity. No matter at what time $t$ we perform our measurement, the measurement result depends not only on the value of the signal $x(t)$ at this moment of time, but also on the values $x(s)$ at nearby moments of time.

In some cases, we are interested in the local behavior of the signal. In this case, we try to measure values which are as close to $x(t)$ as possible. F-transform values are an example of such a local analysis.

In other cases, we are interested in the global trend. In such cases, instead of concentrating on a short-term time interval, we deliberately measure the signal over a long period of time.

**Resulting idea.** To most adequately reconstruct the signal, we should be able to adequately reproduce both its local and its global characteristics. By definition, F-transform adequately represents the local characteristics, no matter what membership function $A(t)$ we select. So, it is reasonable to select a membership function which most adequately represents global characteristics.

Let us describe this idea in precise terms.

## 3. Which Global Characteristics Should We Represent: Discussion

**Need for linearization.** Signals are usually weak. Thus, for any quantity $q$ that depends on this signal $x(t)$ – be it local or global – we should be able to ignore terms which are quadratic or higher order in terms of $x(t)$ and thus retains only the linear terms in the corresponding dependence. As a result, we should only consider linear quantities, i.e., quantities of the type $q = \int q(t) \cdot x(t)\,dt$.

**Which linear quantities should we select?** Of course, when we perform F-transform, we lose some information about the signal. Indeed, on each time interval, we replace infinitely many values $x(t)$ corresponding to infinitely many moments of time $t$ from this interval, with finite many values of the corresponding F-transform. Thus, we cannot perfectly reconstruct all possible global characteristics $q$ – since from the values of all these characteristics, e.g., of the integrals $\int_{-\infty}^{t} x(s)\,ds$ – we would be able to uniquely reconstruct all the values $x(t)$.

Thus, we need to select the most appropriate global characteristics.

**How to define what is most appropriate?** In different situations, different global characteristics may be more appropriate. In this paper, instead of trying to list specific notions of appropriateness, we will consider all possible criteria of this type.

Interestingly, it turns out that all reasonable criteria of this type lead, in effect, to the same family of optimal global characteristics – and the only way to reconstruct these characteristics exactly is to use triangular membership functions.

Let us describe all this in precise terms.

## 4. Towards Precise Formulation of the Problem

**Towards describing what is more appropriate and what is less appropriate.** As we have mentioned, all global characteristics have the form $q = \int q(t) \cdot x(t)\,dt$. Thus, selecting a characteristic is equivalent to selecting the corresponding function $q(t)$.

This function $q(t)$ may be discontinuous, as in the above example of a characteristic $\int_{-\infty}^{t} x(s)\,ds$. However, at least it should be measurable (non-measurable functions cannot be defined without using the Axiom of Choice, which means that they are not definable).

Of course, if we can reconstruct the value $\int q(t)\cdot x(t)\,dt$, then, for every real value $c$, we can also reconstruct the related value $\int (c\cdot q(t))\cdot x(t)\,dt$, since this related value is simply equal to

$$c\cdot\int q(t)\cdot x(t)\,dt.$$

Thus, strictly speaking, a characteristics is represented not by a single function, but by the entire family $\{c\cdot q(t)\}_{c\neq 0}$ of the related functions. So, we arrive at the following definition.

**Definition 1.** *By a* characteristic *or, alternatively, a* family, *we mean a family of the type* $\{c\cdot q(t)\}_{c\neq 0}$, *where* $q(t)$ *is a given measurable function, and* $c$ *runs over all possible non-zero real numbers.*

**Discussion.** What do we mean when we say that some characteristic (family) are more appropriate and some are less appropriate? We mean that we have some criterion according to which, for every two families $F$ and $G$, we can say one of the three things:

- we can say that $F$ is more appropriate than $G$; we will denote this by $G \prec F$;
- we can say that $G$ is more appropriate than $F$; we will denote this by $F \prec G$;
- or we can say that the two characteristics are equally appropriateness; we will denote this by

$$F\sim G.$$

No matter what is the criterion, we have these relations. Thus, we can simply make these relations the definition of a criterion.

Of course, we need to make sure that these relations are consistent: e.g., if $F$ is better than $G$ and $G$ is better than $H$, then $F$ should be better than $H$. Thus, we arrive at the following definition.

**Definition 2.** *By a* criterion *for selecting a characteristic, we means a pair of relations* $\langle\prec,\sim\rangle$ *that satisfies the following properties:*

- *for every two characteristics $F$ and $G$, we have one of only one of three options:*

$$F\prec G,\quad G\prec F,\ \text{and}\ F\sim G;$$

- *if $F\prec G$ and $G\prec H$, then $F\prec H$;*
- *if $F\prec G$ and $G\sim H$, then $F\prec H$;*
- *if $F\sim G$ and $G\prec H$, then $F\prec H$;*
- *if $F\sim G$ and $G\sim H$, then $F\sim H$;*
- *$F\sim F$, and*
- *if $F\sim G$, then $G\sim F$.*

**Discussion.** The whole purpose of selecting a criterion is to use this criterion for selecting the best (most adequate) characteristic, i.e., a characteristic which is better – according to this criterion – than any other characteristic.

So, if there is no such optimal characteristics, the corresponding criterion is useless. But what if there are several characteristics which are all the most appropriate according to the given criterion? In this cases, we can use this non-uniqueness to optimize something else. For example, if several characteristics are equally good in terms of accuracy with which we can predict the future behavior of the signal, then we can select among them the characteristic which is the easiest to compute. As a result, we get, in effect, a new criterion, according to which $F$ is better than $G$ if:

- either $F$ better than $G$ according to the original criterion,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2019 doi:10.20944/preprints201904.0253.v1



- or $F$ equivalent to $G$ in terms of the original criterion but better according to the additional criterion.

If for the new criterion, we still have several different optimal characteristics, we can then optimize something else, etc., until we reach a final criterion for which there is exactly one optimal characteristic.

**Definition 3.**

- *We say that a characteristic $F$ is* optimal *with respect to the criterion $\langle \prec, \sim \rangle$ if for every characteristic $G$, we have $G \prec F$ or $G \sim F$.*
- *We say that the criterion is* final *if there exists exactly one characteristic which is optimal with respect to this criterion.*

**Need for scale-invariance.** A signal $x(t)$ describes how the value of a physical quantity $x$ depends on time. We may have a starting point for the corresponding process, which provides a natural starting point for measuring time, but in general, the numerical value of time depends on what unit we use for measuring time. We can use seconds or minutes or hours – the time interval will be the same but the numerical values will change.

When we replace the original unit for measuring time with a new unit which is $\lambda$ times smaller, then all numerical values of time are *re-scaled*, i.e., multiplied by $\lambda$. For example, if we go from seconds to milliseconds, all numerical values are multiplies by 1000. The function $q(t)$ in the new unit becomes $q(\lambda \cdot t)$.

It is reasonable to require that the relative quality of different characteristics should not change if we simply change the unit used for measuring time, without changing anything of substance. In other words, it is reasonable to require that the criterion be "scale-invariant". Here is a precise definition.

**Definition 4.** *We say that a criterion $\langle \prec, \sim \rangle$ is* scale-invariant *if for every two functions $q(t)$ and $r(t)$ and for every $\lambda > 0$, the following two conditions hold:*

- *if $\{c \cdot q(t)\}_c \prec \{c \cdot r(t)\}_c$, then $\{c \cdot q(\lambda \cdot t)\}_c \prec \{c \cdot r(\lambda \cdot t)\}_c$;*
- *if $\{c \cdot q(t)\}_c \sim \{c \cdot r(t)\}_c$, then $\{c \cdot q(\lambda \cdot t)\}_c \sim \{c \cdot r(\lambda \cdot t)\}_c$.*

**Discussion.** We want to find all membership functions that allow us to reconstruct the most adequate global characteristics. To find these functions, we will first describe which characteristics are the most adequate. Then, we will analyze which membership functions allow us to reconstruct the values of these characteristics from the results of the F-transform.

## 5. Which Characteristics Are the Most Adequate: Preliminary Result

**Discussion.** In the previous section, we argued that the most adequate global characteristic must be optimal with respect to some final scale-invariant criterion. Let us describe all such characteristics.

**Proposition 1.** *For every final scale-invariant criterion, each optimal characteristic has the form $\{c \cdot x^\beta\}_c$, for some real value $\beta$.*

**Proof.** Let us denote the scaling transformation that transforms a family $F = \{c \cdot q(t)\}_c$ into a re-scaled family $\{c \cdot q(\lambda \cdot t)\}_c$ by $T_\lambda$. In terms of this notation, scale-invariance means that:

- if $F \prec G$, then $T_\lambda(F) \prec T_\lambda(G)$; and
- if $F \sim G$, then $T_\lambda(F) \sim T_\lambda(G)$.

Let $\langle \prec, \sim \rangle$ be the final scale-invariant criterion. Since this criterion is final, there exists exactly one optimal characteristic $F_{\rm opt}$. Let us prove that this characteristic is scale-invariant, i.e., that $T_\lambda(F_{\rm opt}) = F_{\rm opt}$ for all $\lambda > 0$. (This proof is similar to the one given in [6].)

Indeed, since $F_{\rm opt}$ is optimal, it is better than or equivalent to any other characteristic. In particular, for every $G$, the characteristic $F_{\rm opt}$ is better than or equivalent to $T_{1/\lambda}(G)$:

$$T_{1/\lambda}(G) \prec F_{\rm opt} \text{ or } T_{1/\lambda}(G) \sim F_{\rm opt}.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2019 doi:10.20944/preprints201904.0253.v1



By applying scale-invariance, we conclude that $T_\lambda(T_{1/\lambda}(G)) \prec T_\lambda(F_{\rm opt})$ or $T_\lambda(T_{1/\lambda}(G)) \sim T_\lambda(F_{\rm opt})$. However, one can easily check that $T_\lambda(T_{1/\lambda}(G)) = G$.

Thus, for every characteristic $G$, we have either $G \prec T_\lambda(F_{\rm opt})$ or $G \sim T_\lambda(F_{\rm opt})$. By definition of an optimal characteristic, this means that the characteristic $T_\lambda(F_{\rm opt})$ is optimal. However, for the final criterion, there is only one optimal characteristic, so we conclude that $T_\lambda(F_{\rm opt}) = F_{\rm opt}$. Thus, the optimal characteristic is indeed scale-invariant.

By definition, each characteristic has the form $\{c \cdot q(t)\}_c$. Let us denote the function $q(t)$ corresponding to the optimal characteristic by $q_{\rm opt}(t)$. The fact that the optimal family is scale-invariant means, in particular, that for every $\lambda > 0$, the function $q_{\rm opt}(\lambda \cdot t)$ – which belongs to the re-scaled family $T_\lambda(F_{\rm opt})$ – also belongs to the original family, i.e., has the form $c(\lambda) \cdot q_{\rm opt}(t)$ for some value $c(\lambda)$: $q_{\rm opt}(\lambda \cdot t) = c(\lambda) \cdot q_{\rm opt}(t)$. It is known that the only measurable functions satisfying this functional equation are functions of the type $C \cdot t^\beta$; see, e.g., [1]. The proposition is proven.

**Discussion.** Let us now find out which membership functions can allow us to reconstruct these most adequate characteristics.

## 6. Which Membership Functions Enable Us to Reconstruct the Most Adequate Global Characteristics

**Definition 5.** *We say that for a membership function $A(t)$, it is possible to always reconstruct a global characteristic $\int q(t) \cdot x(t)\,dt$ if for every $t_0$ and $h$, the value of this characteristic can be uniquely determined once we know all the values*

$$x_i = \int A\left(\frac{t - t_i}{h}\right) \cdot x(t)\,dt.$$

**Case of $\beta = 0$.** A particular case of the most adequate global characteristic is the case $\beta = 0$, when $q(t) = \text{const}$ and the corresponding global characteristic is simply the integral $\int x(t)\,dt = 1$. This characteristic can always be reconstructed from the F-transform, since we require that $\sum_i A\left(\frac{t - t_i}{h}\right) = 1$ for all $t$ and thus,

$$\int x(t)\,dt = \sum_i \int A\left(\frac{t - t_i}{h}\right) \cdot x(t)\,dt = \sum_i x_i.$$

**General case.** Thus, we should worry only about the case when $\beta \neq 0$. In this case, we have the following result.

**Proposition 2.** *The only membership function $A(t)$ for which it is possible to always reconstruct a most adequate global characteristic with $\beta \neq 0$ is the triangular membership function – it can reconstruct the characteristic $\int t \cdot x(t)\,dt$ corresponding to $\beta = 1$.*

*Comment.* This result provides the desired global explanation of why triangular membership functions are so efficient in F-transform applications.

**Proof.** Let us assume that for some $\beta \neq 0$, the membership function $A(t)$ enables us to always uniquely reconstruct the corresponding characteristic

$$\int t^\beta \cdot x(t)\,dt.$$

Let us first consider the case when $t_0 = 0$, $h = 1$, and the signal $x(t)$ is equal to 0 everywhere except for the interval $[0, 1]$. Then, only two F-transform values are different from 0:

- the value $x_0 = \int_0^1 A(t) \cdot x(t)\,dt$, and
- the value $x_1 = \int_0^1 A(t - 1) \cdot x(t)\,dt$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2019 doi:10.20944/preprints201904.0253.v1



The fuzzy partition requirement implies that $A(t) + A(t-1) = 1$, so

$$A(t-1) = 1 - A(t).$$

The only way to be able to always reconstruct the value $\int_0^1 t^\beta \cdot x(t)\,dt$ from these two values, no matter how the signal $x(t)$ behaves on the interval $[0,1]$, is to have $t^\beta$ equal to a linear combination of $A(t)$ and $A(t-1) = 1 - A(t)$. Thus, the function $t^\beta$ is a linear combination of $A(t)$ and 1, and hence, $A(t)$ is a linear combination of $t^\beta$ and 1, i.e., $A(t) = a + b \cdot t^\beta$.

For $t = 1$, we must have $A(t) = 0$, so $a + b = 0$ and thus, $A(t) = a \cdot (1 - t^\beta)$. For $t = 0$, we must have $A(0) = 1$, so we have $a = 1$ and $A(t) = 1 - t^\beta$ for $t \in [0,1]$. Correspondingly, for $s \in [-1,0]$, due to $A(t-1) = 1 - A(t)$, we have $A(s) = 1 - A(s+1) = (s+1)^\beta$.

Let us now consider a signal which is different from 0 only on the interval $[1,2]$. For this signal, the desired global characteristic has the form $\int_1^2 t^\beta \cdot x(t)\,dt$, and the only non-zero values of F-transform are $x_1 = \int_1^2 (1 - (t-1)^\beta) \cdot x(t)\,dt$ and $x_2 = \int_1^2 (t-1)^\beta \cdot x(t)\,dt$. Thus, the only way to exactly reconstruct the global characteristic is to have $t^\beta$ to be a linear combination of $1 - (t-1)^\beta$ and $(t-1)^\beta$, i.e., as a linear combination of $(t-1)^\beta$ and 1: $t^\beta = a \cdot (t-1)^\beta + b$.

Let us show that $\beta = 1$. For this, we need to show that cases when $\beta > 1$ and $\beta < 1$ are impossible. Indeed, differentiating both sides by $t$, we get $\beta \cdot t^{\beta-1} = a \cdot \beta \cdot (t-1)^{\beta-1}$. If $\beta > 1$, then for $t = 1$, we get $\beta = 0$, which contradicts the assumption that $\beta > 1$. If $\beta < 1$, then for $t = 1$, we get $\beta = \infty$ – also a contradiction.

Thus, $\beta = 1$, so $A(t) = 1 - |t|$, i.e., we indeed have a triangular membership function. The proposition is proven.

*Comment.* Once we have a triangular membership function, it is easy to combine the F-transform values to get an integral of a linear function. For simplicity, assume that we start with the signal which is 0 for $t < 0$, and that $h = 1$. Then, the values $x(t)$ corresponding to $t \in [0,1]$, affect the value $x_0$, with the weight $1 - t$, and the value $x_1$, with weight $t$. If we take the difference $x_1 - x_0$, this difference corresponds to the weight $2t - 1$ on $[0,1]$ (and the weight $2 - x$ for $x \in [1,2]$).

We can normalize the difference $x_1 - x_0$ to get the coefficient at $t$ on $[0,1]$ to be equal to 1. For the resulting normalized linear combination $\frac{1}{2} \cdot (x_1 - x_0)$, on $[0,1]$, we have the weight $t - \frac{1}{2}$, and on $[1,2]$, the weight $1 - \frac{t}{2}$.

On the interval $[1,2]$, the next F-transform value $x_2$ corresponds to the coefficient $t - 1$ (and 0 before that). Thus, by adding $x_2$ with the appropriate coefficient, we can make sure that the linear combination continues to have $t$ with coefficient 1 on the interval $[1,2]$ as well. For that, we need to add $x_2$ with coefficient $\frac{3}{2}$. Then the resulting linear combination $\frac{1}{2} \cdot (x_1 - x_0) + \frac{3}{2} \cdot x_2$ is equal to $t - \frac{1}{2}$ on the whole interval $[0,2]$.

On $[2,3]$, this combination is equal to $\frac{3}{2} \cdot (3 - t)$. So, to make sure that we get a linear combination which is equal to $t - \frac{1}{2}$ on the interval $[2,3]$ as well, we need to add $x_3$ with coefficient $\frac{5}{2}$, etc. At the end, when we reach the end of the time interval on which the signal is defined, the corresponding linear combination gives us the integral

$$\int \left(t - \frac{1}{2}\right) \cdot x(t)\,dt = \int t \cdot x(t)\,dt - \frac{1}{2} \cdot \int x(t)\,dt.$$

Since, as we have mentioned, we can easily determine the integral $\int x(t)\,dt$ by adding all the values of the F-transform, we can thus indeed determine the value of the desired global characteristic $\int t \cdot x(t)\,dt$.

**Author Contributions:** The authors contributed equally to this work.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 April 2019 doi:10.20944/preprints201904.0253.v1



**Funding:** This work was supported by the Banking University of Ho Chi Minh City, Vietnam, and by the US National Science Foundation via grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

**Acknowledgments:** The authors are thankful to Irina Perfilieva for encouragement and valuable discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Aczél, J.; Dhombres, J. *Functional Equations in Several Variables*, Camridge University Press:Cambridge, UK, 2008.
2. Belohlavek, R.; Dauben, J.W.; Klir, G.J. *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press: New York, 2017.
3. Klir, G.; Yuan, B. *Fuzzy Sets and Fuzzy Logic*, Prentice Hall: Upper Saddle River, New Jersey, 1995.
4. Kosheleva, O.; Kreinovich, V. Why triangular membership functions are often efficient in F-transform applications: relation to probabilistic and interval uncertainty and to Haar wavelets", In: *Proceedings of the 17th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU'2018*, J. Medina et al. (eds.), Cadiz, Spain, June 11–15, 2018.
5. Mendel, J.M. *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer: Cham, Switzerland, 2017.
6. Nguyen, H.T.; Kreinovich, V. *Applications of Continuous Mathematics to Computer Science*, Kluwer: Dordrecht, 1997.
7. Nguyen, H.T.; Walker, C.; Walker, E.A. *A First Course in Fuzzy Logic*, Chapman and Hall/CRC: Boca Raton, Florida, 2019.
8. Novak, V.; Perfilieva, I.; Holcapek, M.; Kreinovich, V. Filtering out high frequencies in time series using F-transform, *Information Sciences* **2014**, 274, 192–209.
9. Novak, V.; Perfilieva, I.; Kreinovich, V. F-transform in the analysis of periodic signals, In: *Proceedings of the 15th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS'2012*, Inuiguchi, M., Kusunoki, Y., Seki, M. (eds.), Osaka, Japan, September 24–27, 2012.
10. Novák, V.; Perfilieva, I.; Močkoř, J. *Mathematical Principles of Fuzzy Logic*, Kluwer: Boston, Dordrecht, 1999.
11. Perfilieva, I. Fuzzy transforms: Theory and applications, *Fuzzy Sets and Systems* **2006**, 157, 993–1023.
12. Perfilieva, I. F-transform, In: *Springer Handbook of Computational Intelligence*, Springer Verlag, 2015, pp. 113–130.
13. Perfilieva, I., Danková, M., Bede, B. Towards a higher degree F-transform", *Fuzzy Sets and Systems* **2011**, 180(1), 3–19.
14. Perfilieva, I.; Kreinovich, V.; Novak, V. F-transform in view of trend extraction, In: *Proceedings of the 15th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty CJS'2012*, Inuiguchi, M., Kusunoki, Y., Seki, M. (eds.), Osaka, Japan, September 24–27, 2012.
15. Zadeh, L.A. Fuzzy sets, *Information and Control* **1965**, 8, 338–353.