Scaling-Invariant Description of Dependence Between Fuzzy Variables: Towards a Fuzzy Version of Copulas

Gerardo Muela and Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
gdmuela@miners.utep.edu, vladik@utep.edu

Christian Servin
Computer Science and Information Technology
Systems Department
El Paso Community College
919 Hunter
El Paso, TX 79915, USA
cservin@gmail.com

***Abstract*—To get a general description of dependence between n fuzzy variables $x_1, \ldots, x_n$, we can use the membership function $\mu(x_1, \ldots, x_n)$ that describes, for each possible tuple of values $(x_1, \ldots, x_n)$ to which extent this tuple is possible.**

There are, however, many ways to elicit these degrees. Different elicitations lead, in general, to different numerical values of these degrees – although, ideally, tuples which have a higher degree of possibility in one scale should have a higher degree in other scales as well. It is therefore desirable to come up with a description of the dependence between fuzzy variables that does not depend on the corresponding procedure and, thus, has the same form in different scales. In this paper, by using an analogy with the notion of copulas in statistics, we come up with such a scaling-invariant description.

Our main idea is to use marginal membership functions

$$\mu_i(x_i) = \max_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n} \mu(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n),$$

and then describe the relationship between the fuzzy variables $x_1, \ldots, x_n$ by a function $r_i(x_1, \ldots, x_n)$ for which, for all the tuples $(x_1, \ldots, x_n)$, we have

$$\mu(x_1, \ldots, x_n) = \mu_i(r_i(x_1, \ldots, x_n)).$$

I. Towards a Scaling-Invariant Description of Dependence Between Fuzzy Variables: Formulation of the Problem

Fuzzy degrees: a brief reminder. In many real-life situations, it is important to incorporate expert knowledge and experience into a computer-based system. Experts are often not 100% confident about their statements, they may use heuristic rules that they know to be sometimes false. Thus, it is important not only to describe the expert statements themselves, but also to describe the expert's degree of confidence in different statements.

Experts usually describe their degree of confidence by using words from a natural language, such as "usually", "most probably", "in almost all cases", etc. However, computers are not very efficient in processing natural language, they are more efficient in doing what they were originally designed for – processing numbers. It is therefore reasonable to describe expert's degrees of confidence by numbers.

These degrees represent intermediate situations between the cases when the expert is absolutely sure that the statement is true and the cases when the expert is absolutely sure that the statement is false. In the computers, "true" is usually represented as 1 , and "false" as 0. Thus, it makes sense to represent intermediate degrees of confidence by numbers from the interval $[0, 1]$. This is one of the main ideas behind fuzzy logic; see, e.g., [1], [3], [5].

Different scalings of fuzzy degrees are possible. There are many ways to assign a numerical degree to a natural-language term. For example, we can ask an expert to mark his/her degree of confidence on a scale from, say, 0 to 10. If the expert marks 7, we take the ratio 7/10 as the desired degree of confidence. Alternatively, we can ask the expert to select between getting a certain small monetary award when his/her statement is true in a random situation versus getting the same award with some probability – thus measuring the expert's subjective probability that his/her statement is true.

In general, different methods lead to different numerical degrees. In all these cases, the more confident the expert is in a statement, the larger the numerical degree of confidence. Thus, ideally, the same degrees of confidence in one numerical scale correspond to the same degree of confidence in a different scale, and a larger degree of confidence in one scale corresponds to a larger degree of confidence in a different scale.

Let us select two such scales. For each number $a \in [0, 1]$ representing the degree of confidence as described in the first scale, let $f(a)$ denote the corresponding degree of confidence in the second scale. Then, for every two numbers a and a', $a < a'$ should imply $f(a) < f(a')$, i.e., the function $f(a)$ should be strictly increasing.

It is desirable to come up with a scale-invariant dependence between two fuzzy variables. Often, a term used by an expert depends on two or more real-valued variables. For example, when a medical doctor says, during an annual check-up, that a person is healthy, this judgment is based on considering several numerical values such as body-mass index, blood pressure, glucose level, cholesterol level, etc.

In such cases, for each combination $(x_1, \ldots, x_n)$ of values of the corresponding quantities, we have a degree $\mu(x_1, \ldots, x_n) \in [0, 1]$ to which the expert believes that the corresponding object satisfies the given property (e.g., that the person is healthy).

The corresponding function $\mu(x_1, \ldots, x_n)$ – which, in fuzzy logic, is called a *membership function* – described the dependence between the fuzzy quantities $x_1, \ldots, x_n$. However, the numerical values of this function change if we use a different scale for measuring degrees of certainty – in general, we go from $\mu(x_1, \ldots, x_n)$ to

$$\mu'(x_1, \ldots, x_n) = f(\mu(x_1, \ldots, x_n)).$$

It is therefore desirable to come up with a scale-invariant way to describing this dependence, i.e., with a way that would not change if we re-scale all the degrees of confidence.

What we do in this paper. In this paper, we proposed such a scale-invariant description.

Our main idea is to use a similar situation in probabilistic uncertainty, where there is a known scale-invariant way to describe dependence known as *copulas*; see, e.g., [2], [4].

II. From Copulas to Scale-Invariant Description of Dependence Between Fuzzy Variables

Copulas: reminder. To use the above idea, let us first recall what is a copula.

To describe a distribution of a random variable X, we can use the cumulative distribution function (cdf)

$$F(x) \stackrel{\text{def}}{=} \text{Prob}(X \le x).$$

Similarly, to describe the joint distribution of two random variables X_1 and X_2, we can use a 2-dimensional cdf

$$F(x_1, x_2) \stackrel{\text{def}}{=} \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2).$$

For each of the variables X_i, $i = 1, 2$, we can also described their *marginals*

$$F_i(x_i) \stackrel{\text{def}}{=} \text{Prob}(X_i \le x_i).$$

Once we know the joint cdf $F(x_1, x_2)$, we can determine both marginals as $F_1(x_1) = F(x_1, +\infty)$ and $F_2(x_2) = F(+\infty, x_2)$.

The joint cdf contains the information about the marginals and about the relation between the two random variables. How can we describe just the information about the dependence between the random variables?

Let us give an example. Suppose that the random variables X_1 and X_2 are independent. Independence means, in particular, that

$$\text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2) = \text{Prob}(X_1 \le x_1) \cdot \text{Prob}(X_2 \le x_2),$$

i.e., that $F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$. So, independence does not mean any specific value of $F(x_1, x_2)$, it just means that once we know the values of the two marginals $F_1(x_1)$ and $F_2(x_2)$, we can compute the value of $F(x_1, x_2)$ by multiplying the values of the two marginals.

In general, the dependence between the two random variables can be described by specifying a function $C(u_1, u_2)$ such that for every x_1 and x_2, we get

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)).$$

This function $C(u_1, u_2)$ is known as a *copula*; see, e.g., [2], [4].

For example, the case when the variables X_1 and X_2 are independent are described by the product copula

$$C(u_1, u_2) = u_1 \cdot u_2.$$

What is the fuzzy analog of a marginal distribution? The main idea of a copula approach is that we describe the joint cumulative distribution function $F(x_1, x_2)$ in terms of the marginal distributions $F_i(x_i)$ that describe the probabilities of each of the variables x_i.

We would like to similarly describe the joint membership function $\mu(x_1, x_2)$ in terms of the "marginal" membership functions $\mu_i(x_i)$ describing the degree of possibility of different values x_i of the corresponding quantities x_i.

What is a natural way to describe such marginals? For each pair (x_1, x_2), the value $\mu(x_1, x_2)$ describes the degree to which this pair of values is possible. Based on this information, how can we describe the degree to which some value x_1 of the first quantity is possible? This value is possible if either $(x_1, 0)$ is possible or (x_1, x_2) is possible, or (x_1, x_2) is possible for some other value x_2.

In fuzzy logic, the degree to which a statement $A \vee B \vee \ldots$ is true is described by applying an appropriate "or"-operation (also known as a *t-conorm*) $f_\vee(a, b)$ to the degrees $d(A)$, $d(B)$, …to which individual statements are true, i.e., as

$$f_\vee(d(A), d(B), \ldots).$$

In principle, there are many different "or"-operations: $f_\vee(a, b) = \max(a, b)$, $f_\vee(a, b) = 1 + a - a \cdot b$, etc. However, in our case, for each value x_1, we have infinitely many possible values x_2 and thus, infinitely many statements "pair (x_1, x_2) is possible" that we need to combine by using an appropriate "or"-operation. If we apply an operation $f_\vee(a, b) = a + b - a \cdot b$ to infinitely many degrees, we get a meaningless value 1: e.g., if we combine N values equal to d, we get $1 - (1 - d)^N$ which tends to 1 as $N \to \infty$. The same is true for most other "or"-operations, except for $f_\vee(a, b) = \max(a, b)$.

If we use max, then the degree $\mu_1(x_1)$ to which the value x_1 is possible is equal to the maximum of the degrees $\mu(x_1, x_2)$ corresponding to this x_1 and all possible values x_2:

$$\mu_1(x_1) = \max_{x_2} \mu(x_1, x_2).$$

Similarly, we can define the degree $\mu_2(x_2)$ to which the value x_2 is possible is equal to the maximum of the degrees $\mu(x_1, x_2)$ corresponding to this x_2 and all possible values x_1:

$$\mu_2(x_2) = \max_{x_1} \mu(x_1, x_2).$$

Similarly, in the multi-D case, for each membership function $\mu(x_1, \ldots, x_n)$ and for each variable $i = 1, \ldots, n$, we can consider a marginal membership function

$$\mu_i(x_i) =$$
$$\max_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n} \mu(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n).$$

III. Let Us Consider a Non-Degenerate Case

1-D case. In the 1-D case, it is reasonable to consider continuous membership functions $\mu(x)$ that attain values 0 and 1 either at some values x or at infinity.

From 1-D case to multi-D case. In the multi-D case, it is reasonable to consider membership functions $\mu(x_1, \ldots, x_n)$ for which all the marginal distributions $\mu_1(x_1)$, ..., $\mu_n(x_n)$ are continuous functions that attains values 0 and 1 at some (maybe infinite) values x_i.

IV. How to Get a Scaling-Invariant Description of Dependence: Main Idea

Reminder: we want scaling-invariance. We want to find a description of the dependence that does not change if we re-scale all the degrees of confidence, i.e., if, for some monotonic function $f(x)$, we replace all the values $\mu(x_1, x_2, \ldots, x_n)$ with the new values

$$\mu'(x_1, x_2, \ldots, x_n) = f(\mu(x_1, x_2, \ldots, x_n)).$$

Main idea: let us follow the copulas. Let us use the main idea behind copulas and use marginal membership functions $\mu_i(x_i)$ for this description.

To apply this idea, let us analyze how marginal membership functions change under re-scaling.

How marginal membership functions change under re-scaling. Since the function $f(x_1, \ldots, x_n)$ is increasing, for each x_i, the re-scaled function

$$x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \to$$
$$\mu'(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) =$$
$$f(\mu(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n))$$

attains its maximum at exactly the same value $x_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ as the original function

$$x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \to$$
$$\mu(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n).$$

Thus, we should have

$$\mu_i'(x_i) =$$
$$\max_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n} \mu(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) =$$
$$\mu'(\ldots, x_{i-1}, x_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), x_{i+1}, \ldots) =$$
$$f(\mu(\ldots, x_{i-1}, x_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), x_{i+1}, \ldots)) =$$
$$f\left(\max_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n} \mu(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)\right) =$$
$$f(\mu_i(x_i)).$$

How to describe dependence: a possibility. By our assumption, for each i, the value $\mu_i(x_i)$ continuously changes from 0 to 1. Thus, for each number $r_i \in [0, 1]$, there exists a value v_i for which $\mu_i(v_i) = r_i$. In particular, for each tuple $(x_1, \ldots, x_n)$, such a number $r_i(x_1, \ldots, x_n)$ exists for $v = \mu(x_1, \ldots, x_n)$;

$$\mu(x_1, \ldots, x_n) = \mu_i(r_i(x_1, \ldots, x_n)).$$

V. Examples

Let us consider several examples of 2-D membership functions.

Example 1. The first such example is a Gaussian membership function $\mu(x_1, x_2) = \exp(-x_1^2 - x_2^2)$. For this function, as one can easily check, $\mu_1(x_1) = \exp(-x_1^2)$ and $\mu_2(x_2) = \exp(-x_2^2)$. Thus, e.g., the above definition of the function $r_1(x_1, x_2)$ takes the form

$$\exp(-x_1^2 - x_2^2) = \exp(-(r_1(x_1, x_2))^2).$$

By taking minus logarithm of both sides of this equation, we get

$$x_1^2 + x_2^2 = (r_1(x_1, x_2))^2,$$

hence

$$r_1(x_1, x_2) = \sqrt{x_1^2 + x_2^2}.$$

Example 2. The second example is the membership function

$$\mu(x_1, x_2) = \frac{1}{1 + x_1^2 + x_2^2}.$$

This membership function is increasing in x_1 when $x_1 \le 0$ and decreasing in x_1 when $x_1 \ge 0$. Similarly, it is increasing in x_2 when $x_2 \le 0$ and decreasing in x_2 when $x_2 \ge 0$. Thus,

$$\mu_1(x_1) = \frac{1}{1 + x_1^2}.$$

In this example, the condition $\mu(x_1, x_2) = \mu_1(r_1(x_1, x_2)$ takes the form

$$\frac{1}{1 + x_1^2 + x_2^2} = \frac{1}{1 + r_1(x_1, x_2)^2}.$$

If we take the inverse of both sides, we get

$$1 + r(x_1, x_2)^2 = 1 + x_1^2 + x_2^2,$$

hence

$$r_1(x_1, x_2) = \sqrt{x_1^2 + x_2^2},$$

similarly to the Gaussian case.

Example 3. As a third example, let us take the membership function $\mu_1(x_1, x_2) = \exp(-|x_1| - |x_2|)$. In this case, as one can easily check, $\mu_1(x_1) = \exp(-|x_1|)$, so the formula for $r_1(x_1, x_2)$ has the form

$$\exp(-|x_1| - |x_2|) = \exp(-r_1(x_1, x_2)).$$

By taking minus logarithms of both sides, we get

$$r_1(x_1, x_2) = |x_1| + |x_2|.$$

In this case the dependence-describing function $r_1(x_1, x_2)$ is different from the Gaussian case.

Example 4. In the above three examples, we had "independent" fuzzy variables in the sense that we had

$$\mu(x_1, x_2) = \mu_1(x_1) \cdot \mu_2(x_2).$$

Let us provide an example in which the relation between the variables is more complicated. Specifically, let us consider a Gaussian membership function

$$\mu(x_1, x_2) = \exp(-x_1^2 - x_1 \cdot x_2 - x_2^2).$$

In this case, for each x_1, the membership function $\mu(x_1, x_2)$ attains its largest value $\mu_1(x_1)$ when the expression

$$x_1^2 + x_1 \cdot x_2 + x_2^2$$

attains the smallest possible value. Differentiating this expression with respect to x_2 and equating the derivative to 0, we conclude that $x_1 + 2x_2 = 0$, i.e., that

$$x_2 = -\frac{x_1}{2}.$$

Substituting this maximizing value $x_2(x_1)$ into the original expression for the original membership function $\mu(x_1, x_2)$, we conclude that

$$\mu_1(x_1) = \mu(x_1, x_2(x_1)) =$$

$$\exp\left(-x_1^2 + x_1 \cdot \frac{x_1}{2} - \left(\frac{x_1}{2}\right)^2\right) =$$

$$\exp\left(-x_1^2 + \frac{1}{2} \cdot x_1^2 - \frac{1}{4} \cdot x_1^2\right) = \exp\left(-\frac{3}{4} \cdot x_1^2\right).$$

Thus, the requirement that $\mu(x_1, x_2) = \mu_1(r_1(x_1, r_2))$ takes the form

$$\exp(-x_1^2 - x_1 \cdot x_2 - x_2^2) = \exp\left(-\frac{3}{4} \cdot (r_1(x_1, x_2))^2\right),$$

hence

$$(r_1(x_1, x_2))^2 = \frac{4}{3} \cdot (x_1^2 + x_1 \cdot x_2 + x_2^2)$$

and

$$r_1(x_1, x_2) = \frac{2 \cdot \sqrt{3}}{3} \cdot \sqrt{x_1^2 + x_1 \cdot x_2 + x_2^2}.$$

VI. The Resulting Description is Indeed Scaling-Invariant: a Proof

Let us prove that the above-defined functions $r_i(x_1, \ldots, x_n)$ are indeed scaling-invariant.

Indeed, we define the function $r_i(x_1, \ldots, x_n)$ as the function that satisfies the formula

$$\mu(x_1, \ldots, x_n) = \mu_i(r_i(x_1, \ldots, x_n)).$$

If we re-scale membership values, i.e., replace $\mu(x_1, \ldots, x_n)$ with

$$\mu'(x_1, \ldots, x_n) = f(\mu(x_1, \ldots, x_n)),$$

and $\mu_i(x_i)$ with $\mu'_i(x_i) = f(\mu_i(x_i))$, then, by applying the function $f(x)$ to both sides of the above equality, we get the same equality for re-scaled membership degrees:

$$\mu'(x_1, \ldots, x_n) = f(\mu(x_1, \ldots, x_n)) =$$

$$f(\mu_i(r_i(x_1, \ldots, x_n))) = \mu'_i(r_i(x_1, \ldots, x_n)),$$

hence

$$\mu'(x_1, \ldots, x_n) = \mu'_i(r_i(x_1, \ldots, x_n)).$$

So, for the re-scaled membership degrees, we have the exact same functions $r_i(x_1, \ldots, x_n)$.

This means that these functions are indeed scaling-invariant.

VII. Auxiliary Question: What Is the Relation Between Functions $r_i(x_1, \ldots, x_n)$ Corresponding to Different i?

Question. In the above text, we described the dependence between n fuzzy variables $x_1, \ldots, x_n$ by a function $r_i(x_1, \ldots, x_n)$ corresponding to some $i = 1, \ldots, n$. For different indices i, we have different functions $r_i(x_1, \ldots, x_n)$.

What is the relation between functions $r_i(x_1, \ldots, x_n)$ corresponding to different values i? For example, if we know a function $r_i(x_1, \ldots, x_n)$ corresponding to one index i, can we use this function to construct a function $r_j(x_1, \ldots, x_n)$ corresponding to a different index $j \neq i$?

What we prove. In this section, we will prove that such a reconstruction is indeed possible in situations in which:

- all marginal membership functions $\mu_i(x_i)$ are *strictly monotonic fuzzy numbers*, i.e., strictly increase from 0 to 1 when x_i is smaller than some threshold c_i and strictly decrease from 1 to 0 when $x_i \geq c_i$, and
- as values of $r_i(x_1, \ldots, x_n)$, we only select values which are larger than or equal to c_i.

How to reconstruct $r_j(x_1, \ldots, x_n)$ from $r_i(x_1, \ldots, x_n)$: analysis of the problem. By definition of the function $r_i(x_1, \ldots, x_n)$, we have

$$\mu(x_1, \ldots, x_n) = \mu_i(r_i(x_1, \ldots, x_n)).$$

Thus,

$$\mu_j(x_j) =$$

$$\max_{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n} \mu(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n) =$$

$$\max_{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n} \mu_i(r_i(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n)).$$

We have assumed that all the values $r_i(x_1,\ldots,x_n)$ are larger than or equal to threshold values c_i, and that for values $x_i \geq c_i$, the function $\mu_i(x_i)$ is strictly decreasing. Thus, the expression

$$\mu_i(r_i(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n))$$

attains its maximum if and only the expression

$$r_i(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n)$$

is the smallest possible. So,

$$\mu_j(x_j) =$$

$$\max_{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n} \mu_i(r_i(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n)) =$$

$$\mu_i(s_{ij}(x_j)),$$

where we denoted

$$s_{ij}(x_j) \stackrel{\text{def}}{=}$$

$$\min_{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n} r_i(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n).$$

For every x_j, we have $\mu_j(x_j) = \mu_i(s_{ij}(x_j))$ and thus, $\mu_i(x_i) = \mu_j(s_{ij}^{-1}(x_i))$, where, as usual, s_{ij}^{-1} denotes the inverse function. Hence, from the definition of the function $r_i(x_1,\ldots,x_n)$, i.e., from the condition

$$\mu(x_1,\ldots,x_n) = \mu_i(r_i(x_1,\ldots,x_n)),$$

we can conclude that

$$\mu(x_1,\ldots,x_n) = \mu_j(s_{ij}^{-1}(r_i(x_1,\ldots,x_n)),$$

and therefore, that

$$\mu(x_1,\ldots,x_n) = \mu_j(r_j(x_1,\ldots,x_n)),$$

where

$$r_j(x_1,\ldots,x_n) = s_{ij}^{-1}(r_i(x_1,\ldots,x_n)).$$

So, we arrive at the following conclusion.

How to reconstruct $r_j(x_1,\ldots,x_n)$ from $r_i(x_1,\ldots,x_n)$: resulting formulas. Once we know the function $r_i(x_1,\ldots,x_n)$, we compute the auxiliary function

$$s_{ij}(x_j) =$$

$$\min_{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n} r_i(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n),$$

and compute its inverse s_{ij}^{-1}.

After that, we compute

$$r_j(x_1,\ldots,x_n) = s_{ij}^{-1}(r_i(x_1,\ldots,x_n)).$$

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award "UTEP and Prudential Actuarial Science Academy and Pipeline Initiative" from Prudential Foundation.

The authors are thankful to Hung T. Nguyen for valuable discussions.

REFERENCES

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

[2] R. B. Nelsen, *An Introduction to Copulas*, Springer, New York, 1999.

[3] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[4] A. Sklar, "Fonctions de répartition á n dimensions et leurs marges", *Publ. Inst. Statist. Univ. Paris*, 1959, Vol. 8, pp. 229–231.

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