Departmental Technical Reports (CS)   Department of Computer Science

1-1-2013

A Symmetry-Based Approach to Selecting Membership Functions and Its Relation to Chemical Kinetics

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

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

Jorge Y. Cabrera
University of Texas at El Paso, jycabrera@miners.utep.edu

Mario Gutierrez
University of Texas at El Paso, mgutierrez19@miners.utep.edu

Thavatchai Ngamsantivong
King Mongkut's Institute of Technology North Bangkok, tvc@kmutnb.ac.th

Follow this and additional works at: http://digitalcommons.utep.edu/cs_techrep

Part of the Computer Sciences Commons

Recommended Citation

Yet another example is temperature – which can be measured in different scaled differing by a starting point, e.g., in degrees Celsius or in degrees Kelvin. Many formulas remain the same no matter what starting point we use [1].

Ideally, membership functions should reflect this symmetry. Often, our knowledge is imprecise ("fuzzy"). To describe and process such knowledge, L. Zadeh invented the ideas of fuzzy sets; see, e.g., [2], [4], [7]. A fuzzy set on a universal set X is characterized by its membership function $\mu : X \to [0, 1]$.

For shift-invariant quantities, it is desirable that our selection of the corresponding membership functions $\mu : \mathbb{R} \to [0, 1]$ reflect the corresponding shift-invariance.

Comments.

- Symmetries like shift-invariance are often very useful in the analysis of uncertainty, including fuzzy and neural approaches; see, e.g., [3] and references therein.
- In principle, membership functions can be non-smooth and even discontinuous. However, to simplify our analysis, we will assume that all the membership functions are smooth (i.e., differentiable).

First attempt to describe shift-invariance of membership functions. Let us analyze how we can describe this invariance in precise terms. A seemingly natural idea is to require that each membership function is shift-invariant, i.e., that

$$\mu(q) = \mu(q + q_0)$$

for all q and q_0.

Unfortunately, this simply idea does not work: if we impose the above condition, then, by selecting $q_0 = -q$, we get $\mu(q) = \mu(0)$ for all q. Thus, the only membership function $\mu(q)$ which satisfies this condition is the constant function – and such functions do not carry any knowledge about the quantity q.

So, we need a different way of describing shift-invariance of membership functions.

Our idea: first part. Since we cannot require that a *single* membership function is shift-invariant, it is reasonable to require that a collection of *several* membership functions is shift-invariant.

The idea of multiple membership functions is natural in applications of fuzzy techniques: we usually have several rules containing different membership functions $\mu_i(q)$ [2], [4]. If we want to predict the values of a quantity q, then, in some versions of fuzzy system modeling and fuzzy control, we first generate an appropriate linear combination $\sum c_i \cdot \mu_i(q)$ of these membership functions.

Thus, since we cannot require that each membership function $\mu_i(q)$ is shift-invariant, we can require that the set of all such linear combinations is shift-invariant.

Comment. From the mathematical viewpoint, this set of all linear combinations is closed under linear combination and thus, forms a *linear space* of functions. Thus, we arrive at the following definition.

Definition 1.

- *By a finite-dimensional linear space of functions (or simply linear space, for short), we mean the class of all functions of the type $\sum_{i=1}^{n} c_i \cdot \mu_i(q)$, where:*
  - *$n \geq 1$,*
  - *differentiable functions $\mu_1(q)$, ..., $\mu_n(x)$ are fixed (and assumed to be linearly independent), and*
  - *the coefficients $c_1, \ldots, c_n$ can take any real values.*
- *We say that a linear space L is shift-invariant if for every function $f(q)$ from the space L and for every real number q_0, the function $f(q + q_0)$ also belongs to the class L.*

Our idea: second part. The main objective of this paper is to describe all fuzzy-related shift-invariant linear spaces of functions.

This description can be simplified if we take into account that if we have two disjoint linear spaces L_1 and L_2 each of which is shift-invariant, then the set L of linear combinations of all functions from L_1 and L_2 is also shift-invariant. Thus, to describe all shift-invariant families, it is sufficient to describe all *basic* families, i.e., all linear spaces which cannot be decomposed into smaller spaces L_1 and L_2.

Definition 2.

- *If L_1 and L_2 are two linear spaces, then their linear envelope is a space of all functions of the type $f_1(q) + f_2(q)$, where $f_1(q) \in L_1$ and $f_2(q) \in L_2$.*
- *We say that a shift-invariant linear space L is basic if it cannot be represented as a linear envelope of two shift-invariant linear spaces L_1 and L_2.*

Our idea: third part. We are not just interested in general functions, we are interested in membership functions, i.e., in functions whose values are from the interval $[0, 1]$. Let us therefore make the following additional requirement that these values are from the interval $[0, 1]$ – at least when q is non-negative.

We also want to exclude the trivial membership function $\mu(q) \equiv 1$ for all q.

Definition 3. *We say that a linear space of functions L is fuzzy-related if the following two conditions hold:*

- *L is the set of all linear combinations of functions $\mu_1(q)$, ..., $\mu_n(q)$ for each of which $\mu_i(q) \in [0, 1]$ for all $q \geq 0$.*
- *L does not include the constant functions $f(q) \equiv 1$ for all q.*

II. SYMMETRY-BASED APPROACH TO SELECTING MEMBERSHIP FUNCTIONS: MAIN RESULT

Proposition. *Each basic shift-invariant fuzzy-related linear space L is a linear combination of functions*

$$\mu_1(q) = \exp(-\lambda \cdot q), \mu_2(q) = q \cdot \exp(-\lambda \cdot q), \ldots,$$

$$\mu_i(q) = q^{i-1} \cdot \exp(-\lambda \cdot q), \ldots, \mu_n(q) = q^{n-1} \cdot \exp(-\lambda \cdot q),$$

for some $\lambda > 0$.

Proof.

$1°$. Let $L = \{\sum c_i \cdot \mu_i(q)\}$ be a shift-invariant linear space. By definition of shift-invariance, this mean that for each function $f(q)$ from this space (in particular, for each function $f(q) = \mu_i(q)$), the shifted function $f(q + q_0)$ also belongs to this space. For $f(q) = \mu_i(q)$, this means that the function $\mu_i(q + q_0)$ also belongs to the space L. By definition of the space L, this means that the shifted function $\mu_i(q + q_0)$ is a linear combination of the original functions $\mu_1(q), \ldots, \mu_n(q)$, i.e., that

$$\mu_i(q + q_0) = \sum_{j=1}^{n} c_{ij}(q_0) \cdot \mu_j(q), \qquad (1)$$

for some real numbers values $c_{ij}(q_0)$.

The equality (1) holds for all $i = 1, \ldots, n$, so we arrive at the following system of equalities:

$$\mu_1(q + q_0) = \sum_{j=1}^{n} c_{1j}(q_0) \cdot \mu_j(q),$$

$$\cdots$$

$$\mu_i(q + q_0) = \sum_{j=1}^{n} c_{ij}(q_0) \cdot \mu_j(q), \qquad (2)$$

$$\cdots$$

$$\mu_n(q + q_0) = \sum_{j=1}^{n} c_{nj}(q_0) \cdot \mu_j(q).$$

$2°$. By Definition 1, each function $\mu_i(q)$ is differentiable. Let us use the formula (1) to prove that the functions $c_{ij}(q_0)$ are also differentiable. Indeed, let us select n different values $q_1, \ldots, q_k, \ldots, q_n$ of the quantity q, and let us repeat the

formula (1) for each of these values. We then get the following system of equalities:

$$\mu_i(q_1 + q_0) = \sum_{j=1}^{n} c_{ij}(q_0) \cdot \mu_j(q_1),$$
$$\cdots$$
$$\mu_i(q_k + q_0) = \sum_{j=1}^{n} c_{ij}(q_0) \cdot \mu_j(q_k), \qquad (3)$$
$$\cdots$$
$$\mu_i(q_n + q_0) = \sum_{j=1}^{n} c_{ij}(q_0) \cdot \mu_j(q_n).$$

The system (3) is a linear system of equations with n unknowns $c_{i1}(q_0)$, ..., $c_{in}(q_0)$. It is known that in general, each element of the solution to a system of linear equations can be described – via the so-called *Cramer rule* – as a ratio of two determinants, i.e., as a smooth function of the coefficients and the free terms. In our case, the coefficients $\mu_j(q_k)$ are constants (hence, are differentiable), and the free terms $\mu_i(q_k + q_0)$ are also differentiable functions of q_0. Thus, each element $c_{ij}(q_0)$ is a result of applying a smooth function to smooth functions and is, therefore, a differentiable function of q_0.

$3°$. Now that we know that both the functions $\mu_i(q)$ and the functions $c_{ij}(q)$ are differentiable, we can differentiate both sides of the equations (2) and set q_0 to 0. As a result, we get the following system of differential equations:

$$\mu_1'(q) = \sum_{j=1}^{n} C_{1j} \cdot \mu_j(q),$$
$$\cdots$$
$$\mu_i'(q) = \sum_{j=1}^{n} C_{ij} \cdot \mu_j(q), \qquad (4)$$
$$\cdots$$
$$\mu_n'(q) = \sum_{j=1}^{n} C_{nj} \cdot \mu_n(q),$$

where we denoted $C_{ij} \stackrel{\text{def}}{=} c_{ij}'(0)$. Thus, for the functions $\mu_1(q)$, ..., $\mu_n(q)$, we have a system of linear differential equations with constant coefficients. Solutions to such systems are well-known: they have the form $x^k \cdot \exp(-\lambda \cdot q)$, where $-\lambda$ is an eigenvalue of the matrix C_{ij}, and k is an integer corresponding to degenerate eigenvalues, i.e., eigenvalues for which the linear space of the corresponding eigenvectors is more than 1-dimensional:

- if we have only one linear independent eigenvector corresponding to the eigenvalue λ, we only get the term corresponding to $k = 0$;
- if we have two linear independent eigenvectors corresponding to the eigenvalue λ, we get terms corresponding to $k = 0$ and $k = 1$;
- if we have three linear independent eigenvectors corresponding to the eigenvalue λ, we get terms corresponding to $k = 0$, $k = 1$, and $k = 2$;

- $\cdots$

$4°$. Terms corresponding to the same eigenvalue λ form a shift-invariant linear subspace; thus, from the fact that the linear space L is basic, it follows that all the functions from this space correspond to the same eigenvalue λ.

$5°$. In general, for a linear system of differential equations with constant coefficients, we can have positive, zero, and complex eigenvalues, corresponding to negative, zero, or complex values λ.

$5.1°$. In our cases, negative values λ are not possible, since then we will have $\mu_i(q) = q^{i-1} \cdot \exp(-\lambda \cdot q)$ tend to infinity for $q \to \infty$, which contradicts to our assumption that the linear space is fuzzy-related, i.e., that we can select a basis of functions whose values are, for all $q \geq 0$, bounded to the interval $[0, 1]$.

$5.2°$. Zero values λ are also not possible:
- for $i = 1$, $\mu_i(q) = q^{i-1} \cdot \exp(-\lambda \cdot q)$ with $\lambda = 0$ will be a constant function, which contradicts Definition 3, and
- for $i > 1$, this expression tend to infinity for $q \to \infty$, which contradicts to our assumption that the linear space is fuzzy-related, i.e., that we can select a basis of functions whose values are, for all $q \geq 0$, bounded to the interval $[0, 1]$.

$5.3°$. Similarly, complex values $\lambda = a + \mathrm{i} \cdot b$ are impossible, since then terms $\mu_i(q)$ are then proportional to

$$q^{i-1} \cdot \exp(-a \cdot q) \cdot \sin(b \cdot q + \varphi)$$

for some φ, and thus, cannot be non-negative for all $q \geq 0$.

$5.4°$. Since negative, zero, or complex values λ are not possible, we conclude that the value λ must be positive.

The proposition is proven.

III. Meaning of the Above Result and Its Relation to Chemical Kinetics

What is chemical kinetics: brief reminder. Chemical kinetics describes how the concentration of different chemical substances change when the chemical reactions occur. It is usually assumed that the rate of a chemical reaction is proportional to the production of the concentrations of all the substances which are needed for the reaction to take place. For example, for a reaction

$$A + B \to C,$$

the reaction rate is proportional to the product $a \cdot b$ of the concentrations of the substances A and B. Due to this reaction rate $k \cdot a \cdot b$:

- the amounts a and b of substances A and B decrease with this rate, while
- the amount c of the substance C increases with this rate,

i.e., we have

$$\frac{da}{dt} = -k \cdot a \cdot b,$$

$$\frac{db}{dt} = -k \cdot a \cdot b, \qquad (5)$$

$$\frac{dc}{dt} = k \cdot a \cdot b.$$

If we have several different chemical reactions, then, to describe the resulting rate of change of each concentration, we add the rates corresponding to different reactions.

Relation between membership functions and chemical kinetics: an intuitive idea. Let us consider a simple case when we have membership functions corresponding to "small", "medium", and "large". When the value of the quantity is small, we are sure that this value is small and not medium or large. Thus, for this starting value, the membership function corresponding to "small" has a value $s = 1$, while the other two membership functions corresponding to "medium" and "large" have values $m = 0$ and $\ell = 0$.

As we increase q, what was originally small starts slowly transforming into medium, then what was originally medium starts slowly transforming into large, etc. Intuitively, on the quantitative level, this can be described by the following chemical reaction:

$$s \to m;$$

$$m \to \ell. \qquad (6)$$

If we assume that both reactions have the same rate k, then we conclude that for this system of reactions, the corresponding system of equations has the form

$$\frac{ds}{dq} = -k \cdot s;$$

$$\frac{dm}{dq} = k \cdot s - k \cdot m; \qquad (7)$$

$$\frac{d\ell}{dq} = k \cdot m.$$

(Note that since we increase q, not time, the derivatives are with respect to q.)

In view of the above intuitive idea, it is therefore reasonable to interpret the "concentrations" $s(q)$, $m(q)$, and $\ell(q)$ as the degrees to which we believe that the value q is, correspondingly, small, medium, or large. In other words, it is reasonable to interpret the dependencies $s(q)$, $m(q)$, and $\ell(q)$ as the membership functions corresponding to "small", "medium", and "large". Here:

- the quantity q is an analogue of time, and
- the change of the membership degrees $s(q)$, $m(q)$, an $\ell(q)$ with q is analogous to the change of concentrations $a(t)$, $b(t)$, and $c(t)$ of the corresponding chemical reactions with time.

Comment. The above system of reactions is well-studied in chemical kinetics; explicit solutions are given, e.g., in [5], [6].

From an intuitive idea to a formal connection between chemical kinetics and membership functions. Instead of three consequent membership functions corresponding to "small", "medium", and "large", we can consider an arbitrary number n of such functions $\mu_1(q), \ldots, \mu_n(q)$. In this case, the chemical reactions have the form

$$\mu_1 \to \mu_2,$$

$$\mu_2 \to \mu_3,$$

$$\cdots,$$

$$\mu_{n-1} \to \mu_n. \qquad (8)$$
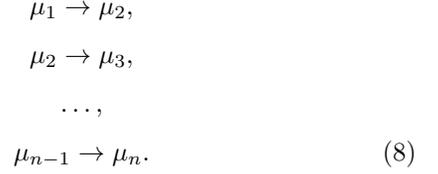
If we assume that all these reactions have the same rate λ, then the corresponding chemical kinetic equations have the form:

$$\frac{d\mu_1(q)}{dq} = -\lambda \cdot \mu_1(q),$$

$$\frac{d\mu_2(q)}{dq} = \lambda \cdot \mu_1(q) - \lambda \cdot \mu_2(q),$$

$$\cdots$$

$$\frac{d\mu_i(q)}{dq} = \lambda \cdot \mu_{i-1}(q) - \lambda \cdot \mu_i(q), \qquad (9)$$

$$\cdots$$

$$\frac{d\mu_{n-1}(q)}{dq} = \lambda \cdot \mu_{n-2}(q) - \lambda \cdot \mu_{n-1}(q),$$

$$\frac{d\mu_n(q)}{dq} = \lambda \cdot \mu_{n-1}(q).$$

Once we set the initial values $\mu_1(0) = 1$ and $\mu_2(0) = \ldots = \mu_n(0) = 0$, this system allows us to uniquely determine the values $\mu_i(q)$ for all $q \geq 0$.

One can easily check that the corresponding solution has the form

$$\mu_i(q) = \frac{\lambda^{i-1}}{(i-1)!} \cdot q^{i-1} \cdot \exp(-\lambda \cdot q). \qquad (10)$$

Indeed, if we plug in this expression into the both sides of the equation (9), we conclude that:

- the left-hand side of the equation (9) has the form

$$\frac{d\mu_i(q)}{dq} = \frac{\lambda^{i-1}}{(i-1)!} \cdot (i-1) \cdot q^{i-2} \cdot \exp(-\lambda \cdot q) -$$

$$\lambda \cdot \frac{\lambda^{i-1}}{(i-1)!} \cdot q^{i-1} \cdot \exp(-\lambda \cdot q) =$$

$$\lambda \cdot \frac{\lambda^{i-2}}{(i-2)!} \cdot q^{i-2} \cdot \exp(-\lambda \cdot q) -$$

$$\lambda \cdot \frac{\lambda^{i-1}}{(i-1)!} \cdot q^{i-1} \cdot \exp(-\lambda \cdot q),$$

- while the right-hand side of the equation (9) has the form

$$\lambda \cdot \mu_{i-1}(q) - \lambda \cdot \mu_i(q) =$$

$$\lambda \cdot \frac{\lambda^{i-2}}{(i-2)!} \cdot q^{i-2} \cdot \exp(-\lambda \cdot q) -$$

$$\lambda \cdot \frac{\lambda^{i-1}}{(i-1)!} \cdot q^{i-1} \cdot \exp(-\lambda \cdot q).$$

So, we indeed have equality for all q.

Comment about normalization. It should be noticed that the corresponding functions (10) are not normalized, in the sense that the maximum of each such function is not equal to 1. If we want normalized membership functions, we must multiply each such function by a corresponding normalizing factor.

To find such a factor, let us find the maximum of the function (10). According to calculus, when a function $\mu_i(q)$ attains its maximum, its derivative $\dfrac{d\mu_i(q)}{dq}$ is equal to 0 (it should be mentioned that the derivative may be equal to 0 at other points as well). Using the formula (9) for this derivative, we conclude that when the maximum is attained, we get $\lambda \cdot \mu_{i-1}(q) - \lambda \cdot \mu_i(q) = 0$, i.e., when

$$\mu_{i-1}(q) = \mu_i(q).$$

Substituting the expressions (10) into both sides of this equality, we conclude that

$$\frac{\lambda^{i-2}}{(i-2)!} \cdot q^{i-2} \cdot \exp(-\lambda \cdot q) = \frac{\lambda^{i-1}}{(i-1)!} \cdot q^{i-1} \cdot \exp(-\lambda \cdot q).$$

Dividing both sides of this equality by the left-hand side, we conclude that $1 = \dfrac{\lambda}{i-1} \cdot q$, so, that at the point where the maximum is attained, we get $q = (i-1) \cdot \dfrac{1}{\lambda}$. It is easy to check that for this value q, the function $\mu_i(q)$ indeed attains its maximum. (Hence, the maxima are equally spaced.)

For this maximizing value q, we have $\lambda \cdot q = i - 1$ and thus, the corresponding value of $\mu_i(q)$ is equal to

$$m_i = \max_q \mu_i(q) = \frac{(i-1)^{i-1}}{(i-1)!} \cdot \exp(-(i-1)). \qquad (11)$$

Hence, the normalized membership function $\widetilde{\mu}_i(q) = \dfrac{\mu_i(q)}{m_i}$, for which $\max_q \widetilde{\mu}_i(q) = 1$, has the form

$$\widetilde{\mu}_i(q) = \frac{\lambda^{i-1}}{(i-1)^{i-1}} \cdot q^{i-1} \cdot \exp((i-1) - \lambda \cdot q). \qquad (12)$$

Comment. The expression (11) for the normalizing coefficient is somewhat complex. We can get a simpler expression for large i if we we use the known asymptotic for the factorial

$$n! \sim \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi \cdot n}.$$

This enables us to conclude that asymptotically,

$$m_i = \max_q \mu_i(q) = \frac{(i-1)^{i-1}}{(i-1)!} \cdot \exp(-(i-1)) \sim$$

$$\frac{(i-1)^{i-1} \cdot e^{i-1}}{(i-1)^{i-1} \cdot \sqrt{2\pi \cdot (i-1)}} \cdot \exp(-(i-1)) =$$

$$\frac{1}{\sqrt{2\pi \cdot (i-1)}}. \qquad (13)$$

Discussion. Normalization does not change the linear space generated by the membership functions.

Thus, we conclude the this section's analogy between chemical kinetics and fuzzy logic leads to the same class of membership functions as the symmetry-based approach: namely, to the class of functions with equally spaced maxima. This provides us with an additional confirmation that these membership functions are reasonable to use.

References

[1] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

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

[3] H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.

[4] H. T. Nguyen and E. A. Walker, *First Course In Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

[5] G. S. Yablonsky, D. Constales, and G. B. Marin, "Coincidences in chemical kinetics: surprising news about simple reactions", *Chemical Engineering Science*, 2010, Vol. 65, pp. 6065–6076.

[6] G. S. Yablonsky, D. Constales, and G. B. Marin, "Equilibrium relationships for non-equilibrium chemical dependencies", *Chemical Engineering Science*, 2011, Vol. 66, pp. 111–114.

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