

Quest for Rigorous Combining Probabilistic and Fuzzy Logic Approaches for Computing with Words

Boris Kovalerchuk

7.1 Introduction

Lotfi Zadeh initiated three fundamental concepts: (1) the concept of a *linguistic variable*, (2) the concept of a *fuzzy set* (with a *membership function* (MF) for linguistic terms that gradually changes between 0 and 1), and (3) the concept a *matrix of linguistic rules* that connect linguistic variables. These concepts are outside of the main stream of concepts used in both the probability theory and the control theory. These concepts are critical for the area that Zadeh later denoted as Computing with Words (CWW) [17]. The *elegance* and *intuitiveness* of these three concepts deeply impressed me when I learned about them a long time ago. This stimulated my interest to contribute to this field. First I noticed from Zadeh's initial work on linguistic variables [15] that operations with fuzzy sets such as min, max, product and others were introduced just as illustrative examples/prototypes of *possible operations* with fuzzy sets. More work was needed to define operations that will be appropriate for CWW. I also noticed that people started to use these "sample" operations without much justification and critical analysis. Next I was impressed by the work of the Zimmermann's team [12], [21] where they analyzed appropriateness of these operations experimentally for linguistic terms metallic and container in German. Later on we conducted a similar experiment [4] in Russian for the same terms, but with different objects. Our results confirmed Zimmermann's negative results in spite of significant differences between these languages in the abilities to create new words by "concatenating" two words.

Using data in ([21], table 6.2) we also computed the average difference of 23.7% between humans' answers for the object to be a metallic container and the $\min(x, y)$ operation value. Here, x is evaluation for the object to be metallic and y to be a container asked separately. In the computation of 23.7% we removed outliers; the difference is greater, 37.76%, if outliers are not removed. For product operation $x \cdot y$ the difference was even higher. Such negative results led to developments of a large collection of other And operations that include compensatory operations [21] that exploited multiple t-norms.

This made fuzzy logic very different from the probability theory with only one And operation, $P(x \& y) = P(y/x) \cdot P(x)$. In fuzzy logic we need to pick up and justify an operation before using it. This is a difficult task and later a way around was found

for fuzzy control. The *conceptual justification of the operation* was substituted by *tuning parameters* of the operation and values $x = m_1(a)$ and $y = m_2(a)$ of membership functions using training data, neural networks and other machine learning methods.

In natural language (NL) CWW getting *training data* is much more difficult especially due to less certainty about the *context* of the NL statements. Therefore, the progress in NL CWW was not as impressive as in fuzzy control.

Lotfi Zadeh posed a set of CWW test tasks and asked whether probability theory can solve these tasks. The recent discussion on relations between fuzzy logic and probability theory for CWW started at the BISC group with a “naïve” question from a student: “*What is the difference between fuzzy and probability?*” Numerous previous debates can be found in the literature (e.g., [1]; [16]). It continued at the uncertainty panel that B. Bouchon-Meunier organized at the World Conference on Soft Computing (WCSC) in May 2011 in San Francisco with L. Zadeh, B. Widrow, J. Kacprzyk, B. Kovalerchuk, and L. Perlovsky as panelists.

7.1.1 Context challenges

Probability Theory, Fuzzy Logic, Dempster-Shafer theory, Rough Sets are oriented to somewhat *different contexts*. However, the appropriate context for a given application often is not clearly formulated, and thus it is very difficult to (*a priori*) select one of the approaches in favor of another. Fuzzy membership functions often are produced (measured) by using frequencies which is a probabilistic way to get MFs. The user of these MFs should get an answer for the question: “Why should we use T-norms and T-conorms with these ‘probabilistic’ MFs instead of probabilistic operations?” The same question is important from the theoretical viewpoint.

7.2 In what sense is probability theory insufficient?

7.2.1 Extreme positions and real challenges

The extreme position known in the fuzzy logic community is expressed by Von Altrock [13]. He stated that *lexical uncertainty* deals with the uncertainty of the *definition of the event* itself and that the probability theory *cannot* be used to model this, as the combination of subjective categories in human decision processes does not follow its axioms. Opposite positions also exist for a long time [e.g., [1]]. Table 7.1 summarizes both extreme views.

Another popular argument that CWW requires a conceptual framework that differs from the probabilistic framework is based on the differences in the nature of stochastic and lexical uncertainties [13]. However, counterexamples exist: water and air have different nature, but hydrodynamics and aerodynamics are modeled by very similar mathematical models.



Fig. 7.1. Lotfi Zadeh at the Computational Intelligence Conference in Honolulu, 08.17.2009. He was the keynote speaker on CWW invited by the author who was the Conference chair.

Zadeh [18], [19] (BISC 03/29/2012) stated *insufficiency of only the standard probability theory* (PT) to deal with CWW. The standard PT is defined by what is found “in textbooks and taught in the classroom,” which is much smaller than the whole scope of the PT. This is an important conceptual difference. From our viewpoint this is not a claim of fundamental insufficiency of PT for CWW. It is a claim that PT models for CWW are not developed without claiming that they cannot be developed within probabilistic framework. We have at least three flavors of probability: (1) frequency-based, (2) subjective, and (2) axiomatic (Kolmogorov’s axioms). The last one abstracts the first two interpretations and there maybe others yet unknown. Thus, thinking about probabilities only as frequencies of repeating events is a very narrow view of probability that should be avoided.

Table 7.1. Comparison of Extreme Probabilistic and Fuzzy Logic positions.

<i>Probabilistic Position</i>	<i>Fuzzy Logic Position</i>
All kinds of uncertainty can be expressed with probability theory.	Stochastic and lexical uncertainties have different nature and require different mathematical models.
Probability theory can model stochastic uncertainty, that a certain event will take place.	Probability theory can model only stochastic uncertainty, that a certain event will take place.
Probability theory can model lexical uncertainty with the uncertainty of the definition of the event itself.	Probability theory cannot model lexical uncertainty with the uncertainty of the definition of the event itself.
Combination of subjective categories in human decision processes does not follow axioms of fuzzy logic theory.	Combination of subjective categories in human decision processes does not follow axioms of probability theory.

To fill deficiencies of the standard PT Zadeh proposed a *Perception-based Probability Theory* and *Generalized Theory of Uncertainty* (GTU) [18], [20], (BISC 03/29/2012) with solutions for CWW including that are consistent with the probabilistic framework. He also stated that he has not attempted to construct an axiomatic approach to GTU, believing that “it will be very hard, perhaps impossible, to do it”. To support this statement Zadeh referenced his *Impossibility Principle*: “The closer you get to reality the more difficult it becomes to reconcile the quest for relevance and applicability with the quest for rigor and precision.”

From my viewpoint the situation is not so hopeless. In several our works it was shown that scientific rigor, relevance, and applicability are reachable when Zadeh’s linguistic variables, membership functions and probabilities combined in what we call a *linguistic context space* [5], [6], [8]. That creates a rigorous base for combination of fuzzy logic and probability concepts. The application of this approach is shown in the next section on one of Zadeh’s test tasks.

Another popular idea is that fuzzy sets and probabilities are *complementary*. I fully agree with this, but for the reasons that differ from just pointing to success of

applications where fuzzy sets and probability combined. The success in application is a reason to take a deeper look to discover the reason of success, but success itself is not sufficient to claim complementarity. We actually took a look at successes of fuzzy control and discovered compelling reasons for success that involves implicit use of probability spaces along with Zadeh's linguistic variables fuzzy sets, and interpolation [7].

7.2.2 Probability vs. Possibility

“What is missing in standard probability theory is the concept of possibility. . . . The absence of this concept *limits* the problem-solving capability of standard probability theory” [20] (04/16/2012, 04/19/2012). Below we attempt to clarify the issue of limitation. Consider a midsize car with five seats.

Question 1: What is the *probability* that 9 people *are* in a midsize car? It is very low, say less than 0.01. To get this answer we watch midsize cars coming to the parking lot and count the number of people in each coming car during the day.

Question 2: What is the *possibility* that 9 people *are* in a midsize car? It is very high, say 0.95. To get this answer we can imagine that four people are sitting on the laps of four others excluding the driver or we can actually seat 9 people as described. This can happen in the case of emergency to be able to escape from a dangerous flood place. Another way to support 0.95 is to notice that Guinness World Record 2011 is 27 people in the 4-seat Mini car. It is obvious if 27 is possible then 9 is easily possible too with much higher possibility.

It seems that this example confirms Zadeh's statement that probability and possibility are different. Now consider another question:

Question 3: What is the *probability* that 9 people *can* be in a midsize car? It is very high, say 0.95. To get 0.95 we can select first randomly 100 people. Next we select 9 people from these 100 people randomly and test if they can sit in the car on laps of each other in the car. Then we repeat this random selection of 9 people multiple times and compute the frequency of success of putting 9 people in the car. Why do we expect that this result will produce a number close to 0.95? We assume that the percent of big people that will have difficulties to sit on the laps in the population is relatively small, say no more than 5%. Next the probability to pick up randomly 9 big people at the same time is small again. More accurate estimates would require knowing the actual share of big people in the whole population.

Now we ask the following questions. Is probability 0.95 as an answer for the Question 3 actually answering about the *possibility* of 9 people in the car? Do we answer Question 2 in this way? Is Question 3 within the probabilistic framework?

It seems that the answers for all these questions are positive, while the standard probability textbooks do not talk about questions like Question 3 as Zadeh pointed out. Note that Question 3 is about *probability of the modal statement* with word

“can”. In the 1980s, P. Cheeseman [1] discussed the probability P on statements that include the word possible: $P(\text{It is possible to put } n \text{ passengers into Carole's car})$. It seems that the probability theory of such modal statements is not developed yet, while there are a few works on modal probability logic. In essence such a new theory would be in the same second-order realm as probabilities on probabilities, fuzzy sets on fuzzy sets, probabilities on possibilities, and possibilities on probabilities. Note that Question 3 is on *probabilities on possibilities* and is formulated in the probabilistic framework that satisfies Kolmogorov's axioms. This conversion of possibilistic Question 2 to Question 3 that is in the probabilistic framework shows that *translation between possibilistic and probabilistic languages is possible*. A particular language can be more convenient, more compact, more intuitive, faster to obtain, etc. However both languages should allow producing the same result with the same rigor. While there is a still active discussion about limits of modeling possibilities by the probability theory it will be good to generate more pro and con examples at this stage of discussion to avoid overgeneralized claims.

7.2.3 Mutual exclusion

There is a popular idea voiced at BICS and multiple publications that concepts like old, young, short, and tall are imprecise overlapping concepts and, therefore need to be modeled by using fuzzy set membership functions not with the probability theory that deals with crisp disjoint (mutually exclusive) elementary events, e.g., die sides. This justification is incomplete and leads to the conceptual difficulties. It does not tell us how to construct these membership functions. A common way to get MF's values is using frequencies of subjective human answers. The fuzzy logic literature is full of such frequencies for computing MFs for ages Young, Old, Middle Age etc, e.g., [2]. This is a probabilistic way to get MFs, which contradicts the idea that PT fundamentally cannot capture such uncertain concepts.

7.2.4 Probability theory and Linguistic uncertainty

Table 7.1 contains a statement from the extreme fuzzy logic position: the probability theory can model *only* stochastic uncertainty that the event will take place, but *cannot* model lexical uncertainty of the definition of the *event*. While PT has an origin in stochastic not linguistic uncertainty as a theory of chances and frequencies back in the 18th century, after A. Kolmogorov published an axiomatic probability theory in 1933, the probability theory moved onto much more abstract level. In this axiomatic theory, elementary events can be elements of *any nature* from sides of dice to *words* such as young and old viewed just as *labels*. Note that word “young” differs from an uncertain real-world concept of being young. It seems that equating a word and an uncertain real-world concept often a reason of the claim that the mutual exclusion axiom of probability prevents modeling such concepts. It can be done via labeling [3].

We also need to have in mind that PT has two very different parts: *abstract* PT as a part of mathematical measure theory and *mathematical statistics* as an area dealing with stochastic uncertainty in the real world. Mathematical statistics matches stochastic uncertainties with abstract PT, but it does not prohibit *matching other linguistic and subjective uncertainties* with the abstract PT. This is done with the development of *subjective PT*, e.g., [14] and works on rigorous combination of *probabilistic concepts with linguistic variables* [3], [6], [8] inspired by a very productive concept of linguistic variables developed by Zadeh [15]. The last approach focuses on *formalizing contexts of lexical uncertainty*.

Zadeh's [18] work in CWW produced generic rules that can be specialized with possibilistic constraints and lead to possibility theory, probabilistic constraints that lead to probability theory; and random-set constraints that lead to the Dempster-Shafer theory of evidence. His perception-based theory of probabilistic reasoning with imprecise probabilities deals with tasks such as: given the perception: Usually Robert returns from work at about 6 p.m.; the question is: What is the probability that he is home at 6:30 p.m.?

7.2.5 Probability and partial truth

Another suggested dividing line between fuzzy logic and probability theory is a statement that PT cannot model partial truth. Zadeh offered the following example for the consideration [18], [20]: "Suppose that Robert is three-quarters German and one-quarter French. If he were characterized as German, the characterization would be *imprecise*, but *not uncertain*. Equivalently, if Robert stated that he is German, his statement would be *partially true*; more specifically, its truth value would be 0.75. Again, 0.75 has no relation to probability."

If we interpret "probability" in last statement as a *common natural language word* then this statement is very consistent with our understanding of this word. However, if we interpret it as a *term of the formal mathematical probability theory* then we may notice that 0.75 can be interpreted as probability because the mathematical term probability has a wider meaning. The axiomatic formal PT is special case of the mathematical measure theory, where 0.75 is just the value of the measure that may have multiple ways to get it. Moreover these ways are outside of the axiomatic theory.

7.3 Linguistic Context Space and multiplicity of solutions of Zadeh's test problems

Zadeh [18], [20] formulated a set of CWW test problems that include the following problems: (1) Usually Robert returns from work at about 6 p.m. What is the probability that he is home at 6:30 p.m.? (2) Probably John is tall. What is the probability that John is short? (3) What is the probability that my car may be stolen? (4) How long does it take to get from the hotel to the airport by taxi?

Zadeh's approach to such problems [18] is (i) representing uncertain concepts listed in the problem using fuzzy sets and/or probabilities, and then (ii) using these representations to the answers. In (1) uncertain concepts are "about" and "usually". In (2) uncertain concepts are "tall", "short", and "probably".

The *linguistic context space* method formally defined in [5]; [6] adds the *context* of uncertain concepts. Adding context is important because context can change the answer. In (2) the context sets up frameworks for the scope of words tall and short. In (2) and many other NL tasks context C is not expressed explicitly while it plays an important role to derive a conclusion. Do we have in mind only tall and short alternatives only or a wider context with more alternatives? In (2) it is intuitively clear that the answer "Probably John is short" is not correct. Thus, other words are needed to express the probability for John to be short. These words could be highly unlikely, more or less unlikely, fifty-fifty, probable, highly probable, or many others. The choice of words can change the answer, but it is not derivable from (2). It depends on the context.

Zadeh [19] (BISC, 8.17.2011) proposed three versions of this problem. "Given: *Probably John is tall*. Version 1: *What is the probability that John is short?* Version 2: *What is the probability that John is very short?* Version 3: *What is the probability that John is not very tall?* What can be assumed is that the imprecise terms tall, short, very short, not very tall and probable are labels of *fuzzy sets* with specified *membership functions*. Alternatively, the terms may be assumed to be labels of specified *probability distributions*. The answer should be a fuzzy probability."

Below we focus on version 1 of problem (2). Consider several sets of linguistic terms (linguistic variables) a part of different contexts:

- Set $PJ1$: {improbable, probable}, or {unlikely, probable};
- Set $PJ2$: {false, unlikely, probable, true};
- Set $PJ3$: {false, possible, probable, true};
- Set $PJ4$: {false, highly unlikely, more or less unlikely, fifty-fifty, probable, highly probable, true};
- Set $H1$: {short, tall};
- Set $H2$: {very short, short, medium, tall, very tall}.

The choice of these sets can change the solution. For context C_1 with sets unlikely, probable and short, tall a common sense answer is: Unlikely John is short (in context C_1). In a variation of this context the answer can be Improbable that John is short (in context C_1). For context C_2 with sets unlikely, probable and very short, short, medium, tall, very tall, a common sense answer is the same: It is unlikely that John is short (in C_2). Extra terms in the linguistic variable for the height did not change the answer while they have expanded the context. For context E_3 with set $PJ4$ that contains terms "highly unlikely" and "more or less unlikely" and set short, tall we have *two alternative answers*: It is highly unlikely that John is short and it is more or less unlikely that John is short (in E_3).

How to get these results computationally? Assume that John's height is 180 cm with the probability $p(\text{tall}, 180) = 0.8$ and the probability that he is short $p(\text{short}, 180) = 0.2$. A voting experiment could give these numbers. Say, 100 people vote whether a man of height 180 cm is tall or short. Alternatively, one person can give these numbers as personal subjective probabilities. In Kolmogorov's terms here we have a probability space with two elementary events $\{\text{short}(180), \text{tall}(180)\}$ and probabilities of these elementary events $p(\text{short}(180)) = 0.8$ and $p(\text{tall}(180)) = 0.2$. Here the mutual exclusion follows from the fact that words short and tall are different.

The probability is defined on *linguistic labels* (short, tall) that are distinct, not on *the natural language concepts of short and tall people that have no sharp border*. For any other height h we can construct similar sets of elementary events $\{\text{short}(h), \text{tall}(h)\}$ with probabilities of elementary events $p(\text{short}(h)) = x$ and $p(\text{tall}(h)) = 1 - x$. This is a very important distinction that we build a Kolmogorov's probability space on labels not fuzzy concepts "short" and "tall". As was already pointed out above the common critic of probability concept in the fuzzy logic community is that probability cannot be defined on the overlapping fuzzy concepts such as "short" and "tall". As we show it is not required to be able to solve this test problem. It is sufficient to build a probability space on a set of labels. Labels as different words are distinct and mutually exclusive.

Figure 7.2 (left) shows these probability spaces for each height h . This figure resembles a set of triangular membership functions commonly used to represent fuzzy linguistic variables. A probability space $S(180)$ is shown as a pair of circles on a vertical line at point $h=180$. As we can see from this figure, *just two membership functions serve as a compact representation of many simple probability spaces* described above. This is a *fundamental representational advantage of Zadeh's fuzzy linguistic variables vs. multiple small probability spaces*. In other words, *few fuzzy membership functions in a linguistic variable provide a quick way to build a huge set of simple probability spaces*. In this sense, fuzzy membership functions and probabilities are complimentary not contradictory. Thus, they are mutually beneficial by combining *fast model development and rigor*. More details are in [6].

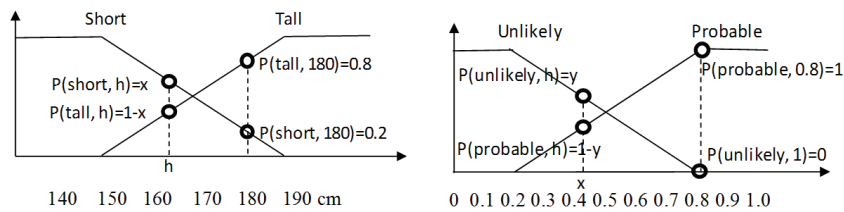


Fig. 7.2. Left: Sets of probability spaces $S(h)$ for elementary events $\{\text{short}(h), \text{tall}(h)\}$ for each height h . Right: Linguistic probabilities *unlikely*, *probable*.

Now we use a set $\{\text{unlikely}, \text{probable}\}$ and build a set of elementary events $\{\text{unlikely}(0.8), \text{probable}(0.8)\}$ for probability value 0.8 with, say, $P(\text{probable}, 0.8) = 1$ and $P(\text{unlikely}, 0.8) = 0$. Other values of probabilities have own sets of elementary events, e.g., $\{\text{unlikely}(0.3), \text{probable}(0.3)\}$ with, say, $P(\text{probable}, 0.3) = 0.25, P(\text{unlikely}, 0.3) = 0.75$.

In general we define a probability space $\{\text{unlikely}(x), \text{probable}(x)\}$ with $P(\text{probable}, x) = 1 - y$ and $P(\text{unlikely}, x) = y$. All of these multiple probability spaces are shown visually in Figure 7.2 (right), and are also *compactly* represented by just two membership functions $\mu_{\text{unlikely}}(h)$ and $\mu_{\text{probable}}(h)$.

Now having $P(\text{tall}, 180) = 0.8$ and $P(\text{probable}, 0.8) = 1$, we convert a numeric $p(\text{tall}, 180) = 0.8$ into a linguistic answer $p(\text{tall}, 180) = \text{probable}$. Formally, it can be done by computing $\max \{P(\text{probable}, 0.8), P(\text{unlikely}, 0.8)\}$ to identify a linguistic term that best fits the 0.8. Similarly having $P(\text{short}, 180) = 0.2$ and $P(\text{unlikely}, 0.2) = 1$ we convert a numeric $p(\text{short}, 180) = 0.2$ into a linguistic $p(\text{short}, 180) = \text{unlikely}$.

For a more general case of a set $\{\text{false}, \text{unlikely}, \text{fifty-fifty}, \text{probable}, \text{true}\}$ and a set $\{\text{short}, \text{tall}\}$ the logic of computations is the same. We will have more probabilities, say, $P(\text{false}, 0.9) = 0, P(\text{unlikely}, 0.9) = 0, P(\text{fifty-fifty}, 0.9) = 0.2, P(\text{probable}, 0.9) = 1, P(\text{true}, 0.9) = 0.9$, but with the same linguistic result of computation: it is unlikely that John is short. However the numeric value of this probability will differ from 0.2 obtained for a smaller set $\{\text{unlikely}, \text{probable}\}$.

7.4 Conclusion and Prospects for Future

As was shown in section 3, just two membership functions serve as a *compact representation* of many simple probability spaces. This is a *fundamental representational advantage of Zadeh's fuzzy linguistic variables vs. multiple small probability spaces*. These few (typically 5-7) fuzzy membership functions within a linguistic variable provide a quick way to build hundreds of simple subjective probability spaces. This is a way how fuzzy membership functions and probabilities become *complementary not contradictory* and mutually beneficial by combining *fast model development and rigor*.

There are also other emerging ways to meet the quest for rigor by developing *more general uncertainty theories* such as (1) by incorporating both rational and irrational agents that generate uncertain statements [11], and (2) by developing an *operation approximation theory* where fuzzy logic T-norm operations are considered as one-dimensional approximations of multidimensional operations in the lattice [9]. More comments on future prospects are in [10].

I believe that the long-term debates between adepts of extreme fuzzy and probabilistic “churches” is in fact a hidden discussion about the *level of acceptance* of the Impossibility Principle quoted above, that Zadeh elegantly formulated. In particular it is the difference in the *level of scientific rigor* and heuristics that are considered as acceptable to get a result relevant to real world challenges. The probability theory itself has got its rigorous foundation only in 1933 with Kolmogorov's axioms after

over 200 years of existence with multiple impressive results and deficiencies. Fuzzy logic is much younger. The history of science has many other examples when new theories reached rigor much later than they produced impressive and useful results along with “results” that were rejected later under more rigorous foundations. Thus, I believe that the productive future in CWW and modeling uncertainty in general is in switching from arguing which extreme position is more wrong to *searching for ways of how to make fuzzy logic and its combination with the probability theory more rigorous, while being still relevant to real world challenges*. To do this we first need to come up to the common concept of what is not rigorous in fuzzy logic. The probability theory would not have gotten a rigorous foundation if (1) its deficiency had not been recognized and (2) the fields of mathematics that it is based on had not been developed before, such as the set theory and the measure theory. I believe that if we follow this constructive path then at some moment the “fight” between the “churches” will be over without any specific effort and the new generalized theories and practice will emerge.

References

1. Cheeseman, Peter: Probabilistic vs. Fuzzy Reasoning. In: Kanal, Laveen N. and John F. Lemmer (eds.): *Uncertainty in Artificial Intelligence*, Amsterdam: Elsevier (North-Holland), 1986, pp. 85–102.
2. Hall, Lawrence O.; Szabo, Sue and Abraham Kandel: On the Derivation of Memberships for Fuzzy Sets in Expert Systems, *Information Sciences*, vol. 40, 1986, pp. 39–52.
3. Hisdal, Ellen: *Logical Structures for Representation of Knowledge and Uncertainty*, New York: Springer (Physica), 1998.
4. Kovalerchuk, Boris and Vladimir Talianski: Comparison of Empirical and Computed Values of Fuzzy Conjunction, *Fuzzy Sets and Systems*, vol. 46, 1992, pp. 49–53.
5. Kovalerchuk, Boris and George J. Klir: Linguistic Context Spaces and Modal Logic for Approximate Reasoning and Fuzzy Probability Comparison. In: Proceedings of the Third International Symposium on Uncertainty Modeling and Analysis and NAFIPS' 95, IEEE Press, A23 A28, 1995.
6. Kovalerchuk, Boris: Context Spaces as Necessary Frames for Correct Approximate Reasoning, *International Journal of General Systems*, vol. 25 (1), 1996, pp. 61–80.
7. Kovalerchuk, Boris: Second Interpolation for Fuzzy Control. In: Fifth IEEE International Conference on Fuzzy Systems. New Orleans, 1996, pp. 150–155.
8. Kovalerchuk, Boris and Evgenii Vityaev: *Data Mining in Finance: Advances in Relational and Hybrid Methods* (chapter 7 on fuzzy systems), Boston: Kluwer, 2000.
9. Kovalerchuk Boris: Interpretable Fuzzy Systems: Analysis of T-norm interpretability, Proceedings of the WCCI World Congress on Computational Intelligence, Barcelona, July 2010, pp. 2035 ff.
10. Kovalerchuk, Boris: BISC-group posts, 2010-2012, <http://mybisc.blogspot.com/>
11. Resconi, Germano and Boris Kovalerchuk: Agents' Model of uncertainty, *Knowledge and Information Systems*, vol. 18 (2), 2009, pp. 213–229.
12. Thole, U.; Zimmermann, Hans-Jürgen and Peter Zysno: On the Suitability of Minimum and Product Operations for the Intersection of Fuzzy Sets, *Fuzzy Sets and Systems* vol. 2, 1979, pp. 173–186.
13. Von Altrock, Constantin: *Fuzzy Logic and Neuro Fuzzy Applications in Business and Finance*, New Jersey: Prentice Hall, 1997.
14. Wright, George and Peter Ayton (eds.): *Subjective Probability*, Chichester, New York: Wiley, 1994.
15. Zadeh, Lotfi A.: The Concept of a Linguistic Variable and its Application to Approximate Reasoning – I, *Information Sciences*, vol. 8, 1977, pp. 199–249.

16. Zadeh, Lotfi A.: Discussion: Probability Theory and Fuzzy Logic are Complementary Rather than Competitive, *Technometrics*, vol. 37 (3), 1996, pp. 271–276.
17. Zadeh, Lotfi A.: From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions, *IEEE Transactions on Circuits Systems*, vol. 45, 1999, pp. 105–119.
18. Zadeh, Lotfi A.: Toward a Perception-based Theory of Probabilistic Reasoning with Imprecise Probabilities, *Journal of Statistical Planning and Inference*, vol. 105, 2002, pp. 233–264.
19. Zadeh, Lotfi A.: A Note on Z-numbers, *Information Sciences*, vol. 181, 2011, pp. 2923–2932.
20. Zadeh, Lotfi A.: BISC-group posts, 8.17.2011, 03/29/2012, 04.16.2012, 04.19.2012 <http://mybisc.blogspot.com/>
21. Zimmermann, Hans-Jürgen: *Fuzzy Sets, Decision Making, and Expert Systems*, Boston: Kluwer-Nijhoff, 1987.