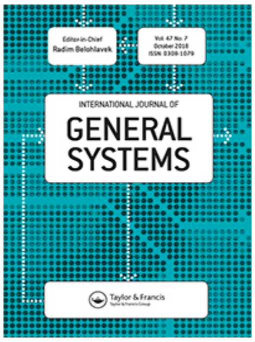

International Journal of General Systems

ISSN: 0308-1079 (Print) 1563-5104 (Online) Journal homepage: http://www.tandfonline.com/loi/ggen20

On Zadeh's problem in probability theory

R. Belohlavek

To cite this article: R. Belohlavek (2018) On Zadeh's problem in probability theory, International Journal of General Systems, 47:7, 648-654, DOI: 10.1080/03081079.2018.1494734

To link to this article: https://doi.org/10.1080/03081079.2018.1494734

Published online: 20 Jul 2018.

INTERNATIONAL JOURNAL OF GENERAL SYSTEMS
2018, VOL. 47, NO. 7, 648–654
https://doi.org/10.1080/03081079.2018.1494734

On Zadeh's problem in probability theory

R. Belohlavek

Department of Computer Science, Palacký University, Olomouc, Czech Republic

ABSTRACT
We pay tribute to Lotfi Zadeh, a long-term member of the editorial board of this journal, by drawing attention to his challenges to classical probability theory. In particular, we consider the following problem: Probably John is tall. What is the probability that John is short? What is the probability that John is very short? What is the probability that John is not very tall? We discuss possible solutions and general significance of this and similar problems.

ARTICLE HISTORY
Received 30 April 2018
Accepted 26 June 2018

KEYWORDS
Probability theory; fuzzy logic

1. Zadeh's challenges to classical probability theory

Fuzzy logic not only represents the most important invention of Lotfi Zadeh but also a lasting subject of his research (Zadeh 2015). Zadeh regarded fuzzy logic as a principal means for modeling semantics of natural languages, which became a central theme of his explorations since the early 1970s (Zadeh 1973, 1975a,b,c, 1994a,b, 2005, 2006); see also Klir and Yuan (1996) and Yager et al. (1987).

Zadeh considered it of utmost importance to develop semantics of natural languages that would enable one to solve problems described in natural language. He repeatedly argued that approaches based on classical set theory and classical probability theory do not have this capability. To support his contention, he used to present examples of problems that commonly arise in human reasoning. These "simple examples", as Zadeh called them, exhibit typical features of natural language expressions, namely various kinds of uncertainty of words such as "small", "very", "most", "likely", "probable", etc.

As an example, consider the following problem, which Zadeh presented as a problem in probability theory:

$$\text{Probably John is tall.} \tag{1}$$

$$\text{What is the probability that John is short?} \tag{2}$$

$$\text{What is the probability that John is very short?} \tag{3}$$

$$\text{What is the probability that John is not very tall?} \tag{4}$$

CONTACT R. Belohlavek radim.belohlavek@acm.org

As Zadeh argued, such problems cannot be solved by classical probability theory, i.e. probability theory based on Kolmogorov's axioms, but they may be approached via concepts derived from the notion of a fuzzy set. Solutions to these problems are, nevertheless, rarely discussed in the literature.

The aim of this paper is to draw attention to such problems and emphasize their significance both from the theoretical as well as applied point of view. In the next section, we outline a possible solution of the problem described above. The significance of Zadeh's problems to challenge classical probability theory is discussed in Section 3.

The problem regarding John's height, described in (1)–(4), has been circulated by Zadeh in 2011. In May 2011, Professor Zadeh addressed me by email with a preliminary version of this problem and asked me to send him a solution. He sent me the present version of the problem later, in August 2011. I came up with a solution and further questions regarding the assumptions of the problem in late August, and then with a final version of my solution in early September. During our email correspondence in early September, it turned out that my solution was almost identical with Zadeh's own solution.[1] Note also that Zadeh presented this problem along with other challenging problems in his talk at Logic Colloquium at Berkeley (Zadeh 2011).

Before presenting a possible solution to the problem regarding John's height, let me note that Zadeh repeatedly emphasized on various occasions that most people do not consider this problem a mathematical problem. According to Zadeh, not only was this problem usually regarded as ill-posed by classical mathematicians, it was not considered as a mathematical problem even by most researchers in fuzzy logic. A particular point we attempt to make is that the problem indeed may be regarded as a mathematical problem.

Note also that many studies on the various kinds of connections between probability theory and fuzzy set theory have been presented in the literature, see Belohlavek, Dauben, and Klir (2017) and, in particular, Dubois and Prade (1993), Nguyen (1997) and Bronevich and Karkishchenko (2002), and that we do not comment on these connections in the present paper.

2. A possible solution of Zadeh's problem regarding John's height

To solve the problem regarding John's height mathematically,[2] we first need to transform the description of condition (1) and the questions (2)–(4), which are provided in natural language, to their appropriate mathematical representation.

Some considerations regarding our assumptions need to be made now. Let us denote the set of possible values of John's height (in cm) by Ω and assume $\Omega = \{150, \dots, 200\}$. Furthermore, denote the set of all probability distributions p_J of John's height on Ω by $\mathcal{D}$.[3] It is convenient for our purpose to rephrase (1) as follows:

$$\text{It is probable that John is tall.} \tag{5}$$

The description of our problem involves vague terms "tall", "short", "very short", "not very tall ", and "probable", whose meaning is naturally described by fuzzy sets.[4] We assume that the above terms are represented by fuzzy sets $T : \Omega \to [0, 1]$, $S : \Omega \to [0, 1]$, $vS : \Omega \to [0, 1]$, $nvT : \Omega \to [0, 1]$, and $Prob : [0, 1] \to [0, 1]$, respectively. That is, $T(\omega)$ is the truth degree in [0, 1] to which a man with height ω is considered tall, and the same with S, vS, and nvT. For $a \in [0, 1]$, $Prob(a)$ is a truth degree in [0, 1] to which an event with probability

a is considered probable. *Prob* satisfies certain natural properties, e.g. is nondecreasing, $Prob(0) = 0$, $Prob(1) = 1$, we also assume that *Prob* is symmetric w.r.t. 0.5 in that

$$Prob(0.5 + a) = 1 - Prob(0.5 - a) \tag{6}$$

for each $0 \leq a \leq 1$.

The principal component of our solution consists in observing that:

Condition (1) *represents a restriction on probability distributions of John's height.*

Classical probability theory is able to calculate the probability of, e.g. "John's height is ≥ 180" but not of "John is tall": While ≥ 180 is an event in Ω in the sense of classical probability theory, i.e. $[180, 200] \subseteq \Omega$, the fuzzy set T representing "tall" *is not an event in the classical sense*. Rather, T *represents a fuzzy event.* Zadeh recognized the importance of the notion of a fuzzy event and that of a probability of fuzzy events at the early stage of development of fuzzy sets (Zadeh 1968). For a probability distribution $p_J \in \mathcal{D}$, the corresponding probability $P_J(T)$ of the fuzzy event T is defined as[5]

$$P_J(T) = \sum_{\omega \in \Omega} p_J(\omega) \cdot T(\omega). \tag{7}$$

Now, every possible distribution $p_J \in \mathcal{D}$ of John's heights induces a degree to which p_J satisfies restriction (1), namely the degree $Prob(P_J(T))$ to which it is probable that John is tall. This way, restriction (1) induces a fuzzy set $R : \mathcal{D} \to [0, 1]$ of probability distributions defined for $p_J \in \mathcal{D}$ by

$$R(p_J) = Prob(P_J(T)). \tag{8}$$

It is therefore natural that:

Answers to questions (2)–(4) *are to be represented by restrictions, too,* namely by restrictions on the possible probabilities that John is short, that John is very short, and that John is not very tall.

Answers to questions (2) *and* (3): Consider an arbitrary fuzzy event A. Given a restriction on probability distributions $p_J \in \mathcal{D}$, one naturally regards a probability $a \in [0, 1]$ as possible if there exists a probability distribution p_J satisfying the restriction for which $P_J(A) = a$. Suppose now that the restriction is represented by a set $R \subseteq \mathcal{D}$ of probability distributions. Then the possible probabilities of A also form a set, $\pi_A \subseteq [0, 1]$, defined as

$$a \in \pi_A \quad \text{iff} \quad \exists p_J \in \mathcal{D} : \ P_J(A) = a \text{ and } p_J \in R.$$

Now, if the restriction is represented by a fuzzy set $R : \mathcal{D} \to [0, 1]$, the possible probabilities also form a fuzzy set $\pi_A : [0, 1] \to [0, 1]$, which is defined in accordance with Zadeh's extension principle as follows:

$$\pi_A(a) = \sup\{R(p_J) \mid p_J \in \mathcal{D} \text{ s.t. } P_J(A) = a\}. \tag{9}$$

The above reasoning may directly be applied to answer our questions. For the restriction R defined by (8) and the fuzzy event S, we obtain the fuzzy set π_S assigning to each probability $a \in [0,1]$ the degree $\pi_S(a)$ to which a is possible probability that John is short:

$$\pi_S(a) = \sup\{Prob(P_J(T)) \mid p_J \in \mathcal{D} \text{ s.t. } P_J(S) = a\}. \tag{10}$$

Observe now that for the fuzzy set nT representing "not tall", which is defined by $nT(\omega) = 1 - T(\omega)$, we have

$$\begin{aligned} P_J(nT) &= \sum_{\omega \in \Omega} p_J(\omega) \cdot nT(\omega) = \sum_{\omega \in \Omega} p_J(\omega) \cdot (1 - T(\omega)) \\ &= \sum_{\omega \in \Omega} p_J(\omega) - \sum_{\omega \in \Omega} p_J(\omega) \cdot T(\omega) = 1 - P_J(T). \end{aligned}$$

In addition, since $S \subseteq nT$, i.e. $S(\omega) \leq nT(\omega)$ for each $\omega \in \Omega$, which corresponds to the fact that "short" implies "not tall", we obtain $1 - P_J(nT) \leq 1 - P_J(S)$. Due to the properties of *Prob* representing "probable", cf. the text around Equation (6), we have $Prob(1 - P_J(nT)) \leq Prob(1 - P_J(S))$ and $Prob(1 - a) = 1 - Prob(a)$. With these observations, we obtain from (11) that π_S satisfies

$$\begin{aligned} \pi_S(a) &= \sup\{Prob(P_J(T)) \mid p_J \in \mathcal{D} \text{ s.t. } P_J(S) = a\} \\ &= \sup\{Prob(1 - P_J(nT)) \mid p_J \in \mathcal{D} \text{ s.t. } P_J(S) = a\} \\ &\leq \sup\{Prob(1 - P_J(S)) \mid p_J \in \mathcal{D} \text{ s.t. } P_J(S) = a\} \\ &= Prob(1 - a) = 1 - Prob(a). \end{aligned}$$

The restriction on the probabilities of "John is short" we derived is therefore represented by the condition $\leq 1 - Prob$, which may be expressed symbolically as

The probability that John is short is $\leq 1 - Prob$.

Interestingly, the restriction may be transformed back to a description in natural language, even though losing certain specificity of the solution we obtained. Namely, since $1 - Prob$, i.e. the fuzzy set *nProb*, represents "not probable" and since being less than or equal to "not probable" still means "not probable", our solution leads to the following conclusion described in natural language:

It is not probable that John is short.

This conclusion follows from condition (1) by intuition but note that we have derived it mathematically.

As for question (3), the reader may easily check that the same reasoning applies to the fuzzy event *vS* representing "very short". In particular, for the degree $\pi_{vS}(a)$ of possibility that the probability of "John is very short" equals a we get again $\pi_{vS}(a) \leq 1 - Prob(a)$, i.e. one may conclude again that it is not probable that John is very short.

Answer to question (4):While the restrictions π_S and π_{vS} on the possible probabilities of John being short and very short may be obtained via the extension principle and relationships to the probability of being tall, such relationships are not apparent for the question related to John being not very tall. Namely, unlike "short" and "very short", "not very tall" does not imply "not tall". Hence, we can only express the degree $\pi_{nvT}(a)$ to which it is possible that the probability of John being not very tall equals a as

$$\pi_{nvT}(a) = \sup\{Prob(P_J(T)) \mid p_J \in \mathcal{D} \text{ s.t. } P_J(nvT) = a\}, \tag{11}$$

i.e. as a solution to an optimization problem.

3. Significance of Zadeh's challenges

Most often than not, descriptions of real-world problems are provided in natural language and often pertain to probabilities. As Zadeh repeatedly argued, classical probability theory is not capable of solving these problems. To support his claims, he used to present deceivingly simple problems, such as the problem examined in this paper, which are beyond the scope of classical probability theory. Zadeh's purpose in presenting the challenging problems was not confined to demonstrating the limits of classical probability theory. He considered it important to develop an extension of classical probability theory with the capability to solve these problems, and argued in his papers that fuzzy logic provides a proper framework for such an extension.

It was a typical response to such problems that they were not considered as mathematical problems. This paper helps to demonstrate that such perception is not warranted—the problems lend themselves to mathematical treatment. In particular, a possible solution is to regard the input conditions of the problems as describing in natural language certain restrictions and to regard the questions of the problems as questions concerning other restrictions. The notions naturally suited for such solutions are derived from fuzzy logic. The answers obtained may be simple, and may even be transformed to natural language descriptions in some cases, but may also be complex and typically involve solutions of optimization problems.

I find it proper to put Zadeh's challenging problems, and more generally his studies pertaining reasoning in natural language, in historical perspective, albeit rather briefly. As argued in Belohlavek, Dauben, and Klir (2017), the main theme in logic until the late 19th century has been human reasoning in a broad sense. Logical explorations included both deductive and inductive reasoning, along with numerous issues extending to philosophy and psychology, but reasoning in natural language had always been the main subject. In the mid-19th century, logic started to focus on deductive reasoning, leaving the subject of inductive reasoning to probability theory. Since the late 19th century and early 20th century, which mark the beginnings of modern formal logic, logic has primarily been regarded as concerned with mathematical reasoning rather than reasoning in natural language. This implied considerable simplification but also restrictions of applicability. From this perspective, Zadeh's works in fuzzy logic may be regarded as an attempt to revive reasoning in natural language as a grand theme of logic. Today, this picture is not properly recognized. I believe that Zadeh's attempt is of fundamental importance and that its significance in broad terms will properly be recognized in the years to come.

Notes

1. Let me emphasize that my goal in this paper is not to claim priority in solving this problem but rather discussing the problem and its possible solutions, and thus pay tribute to Zadeh.

2. Note first that we present our solution in terms as simple as possible, in particular not dependent on the rich network of concepts of Zadeh's computing with words, calculus of restrictions, and other frameworks, to keep the presentation short and easily accessible.
3. That is, $p_J \in \mathcal{D}$ means that p_J assigns a number $p_J(\omega)$ in $[0, 1]$ to every ω in Ω in such a way that $\sum_{\omega \in \Omega} p_J(\omega) = 1$.
4. A fuzzy set (Zadeh 1965) in a universe set U is a mapping $A : U \to [0, 1]$; for $u \in U$, $A(u)$ is interpreted as the membership degree of u in A.
5. Zadeh (1968) defines a fuzzy event in a given probability space with outcomes in $\mathbb{R}^n$ and probability measure P as a Borel measurable fuzzy set A and its probability as the Lebesgue–Stieltjes integral $P(A) = \int_{\mathbb{R}^n} A \, \mathrm{d}P$. An adaptation to our situation leads to (7).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributor

Radim Belohlavek received Ph.D. degree in computer science from the Technical University of Ostrava, Czech Republic, in 1998, Ph.D. degree in mathematics from Palacky University, Olomouc, Czech Republic, in 2001, and D.Sc. degree in informatics and cybernetics from the Academy of Sciences of the Czech Republic in 2008. He is a professor of computer science at Palacky University. Dr. Belohlavek's academic interests are in discrete mathematics, logic, uncertainty and information, and data analysis. He published four books (Kluwer, Springer, MIT Press, Oxford University Press) and over 150 papers in conference proceedings and journals. Dr. Belohlavek is a Senior Member of IEEE (Institute of Electrical and Electronics Engineers), and a Member of ACM (Association for Computing Machinery) and AMS (American Mathematical Society), and is a member of editorial boards of several international journals.

ORCID

R. Belohlavek http://orcid.org/0000-0003-4924-3233

References

Bělohlávek, R., J. W. Dauben, and G. J. Klir. 2017. Fuzzy Logic and Mathematics: A Historical Perspective. New York: Oxford University Press.
Bronevich, A. G., and A. N. Karkishchenko. 2002. "Statistical Classes and Fuzzy Set Theoretical Classification of Probability Distributions." In *Statistical Modeling, Analysis and Management of Fuzzy Data*. Studies in Fuzziness and Soft Computing, edited by Carlo Bertoluzza, María-Ángeles Gil, Dan A. Ralescu, Vol. 87, 173–195. Heidelberg: Physica-verlag.
Dubois, D., and H. Prade. 1993."Fuzzy Sets and Probability: Misunderstandings, Bridges and Gaps." In Proceedings of FUZZ IEEE 1993, San Francisco, CA, USA, Vol. 2, 1059–1068.
Klir, G. J., and B. Yuan. 1996. Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh. Singapore: World Scientific.
Nguyen, H. T. 1997. "*Fuzzy Sets and Probability*." Fuzzy Sets and Systems 90: 129–132.
Yager, R. R., S. Ovchinnikov, R. M. Tong, and H. T. Nguyen, eds. 1987. Fuzzy Sets and Applications: Selected Papers by L. A. Zadeh. New York: Wiley.
Zadeh, L. A. 1965. "*Fuzzy Sets*." Information and Control 8: 338–353.
Zadeh, L. A. 1968. "*Probability Measures of Fuzzy Events*." Journal of Mathematical Analysis and Applications 23: 421–427.

Zadeh, L. A. 1973. "*Outline of a New Approach to the Analysis of Complex Systems and Decision Processes*." IEEE Transactions on Systems, Man, and Cybernetics 3: 28–44.
Zadeh, L. A. 1975a. "Calculus of Fuzzy Restrictions, Fuzzy Sets and their Applications to Cognitive and Decision Processes." In *Proceedings of the U.S.Japan Seminar on Fuzzy Sets and Their Applications*, University of California, Berkeley, CA, July 1–4, 1974, edited by L. A. Zadeh, K. S. Fu, K. Tanaka, M. Shimura, 1–39. New York: Academic Press.
Zadeh, L. A. 1975b. "*Fuzzy Logic and Approximate Reasoning*." Synthese 30: 407–428.
Zadeh, L. A. 1975c. "*The Concept of Linguistic Variable and Its Application to Approximate Reasoning—I, II, III*." Information Sciences 8: 199–249. 8:301–357; 9:43–80.
Zadeh, L. A. 1994a. "*Fuzzy Logic, Neural Networks, and Softcomputing*." Communications of the ACM 37: 77–84.
Zadeh, L. A. 1994b. "*Soft Computing and Fuzzy Logic*." IEEE Software 11: 48–56.
Zadeh, L. A. 2005. "*Toward a Generalized Theory of Uncertainty (GTU)—An Outline*." Information Sciences 172: 1–40.
Zadeh, L. A. 2006. "*Generalized Theory of Uncertainty (GTU): Principal Concepts and Ideas*." Computational Statistics and Data Analysis 51: 15–46.
Zadeh, L. A. 2011. "Can Mathematics Deal with Computational Problems which are Stated in a Natural Language?." Talk presented at Logic Colloquium, University of California at Berkeley, September 30, 2011.
Zadeh, L. A. 2015. "*Fuzzy Logic—A Personal Perspective*." Fuzzy Sets and Systems 281: 4–20.