Information Sciences 267 (2014) 334–336

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A note on similarity-based definitions of possibility and probability ☆

Lotfi A. Zadeh *

Department of EECS, University of California, Berkeley, CA 94720-1776, United States

ARTICLE INFO

Article history:
Received 12 December 2013
Received in revised form 22 January 2014
Accepted 30 January 2014
Available online 7 February 2014

Keywords:
Similarity-based possibility
Similarity-based probability

ABSTRACT

Probability theory is a rigorous theory which rests on a firm foundation. The brilliant successes of probability theory are visible to all. What is problematical relates to the basic question: What is probability? There are many answers but no unanimity.

What is suggested in this note are definitions of possibility and probability which are based on fuzzy logic. More concretely, the definitions are based on the concept of similarity. Similarity-based definition of probability reduces to the traditional frequency-based definition of probability when similarity is exact.

1. Introduction

Probability theory is a formalism for dealing with uncertainty. Mathematically, probability theory is a deep and rigorous theory which rests on a firm foundation. What is problematical is not the theory but its bridges to the real world. In real-world settings, there are many simple questions which are hard to answer for probability theory. What is the meaning of: The probability that Vera is middle-aged is 0.7? What is the probability that Robert is rich? What is the meaning of: It is possible but not probable that Pat is young? What is the meaning of: It is very unlikely that there will be a sharp decline in unemployment within the next few months? Such questions are hard to answer because it is difficult to answer the basic question: What is probability? There are many answers but no unanimity. [1,4,6] Additionally, many traditional definitions in probability theory—based as they are on bivalent logic—are not operational in realistic settings. Example. Definition of stationarity.

In this very brief note, conceptually simple definitions of possibility and probability are suggested. The definitions in question are based on fuzzy logic and, more specifically, involve the concept of similarity [8]. For purposes of exposition, it is convenient to employ a simple example to formulate the definitions of similarity-based possibility and probability.

Consider the question: How old is Vera? Obviously, to answer this question it is necessary to have some information about Vera. It is expedient to represent this information as an information vector:

☆ Research supported in part by ONR N00014-02-1-0294, Omron Grant, Tekes Grant, Azerbaijan Ministry of Communications and Information Technology Grant, Azerbaijan University of Azerbaijan Republic and the BISC Program of UC Berkeley.

* Tel.: +1 510 642 4959.

E-mail address: zadeh@eecs.berkeley.edu

http://dx.doi.org/10.1016/j.ins.2014.01.046
0020-0255/

IV

A_1	…	A_n	*Age*
u_1	…	u_n	*?X*

In this vector, the A_i are age-relevant attributes of Vera with known attribute values, $u_1, \ldots, u_n$. Examples. Ages of parents, ages of children and grandchildren, position, education, etc. The unknown attribute is Age. In addition to knowing the attribute values, we have experience-based information (world knowledge) which may be represented as a dataset of the form:

ID

	A_1		A_n	*Age*	*Sim*
Name$_1$				a_1	s_1
Name$_k$				a_k	s_k
Name$_m$				a_m	s_m

In this information dataset, a_j is the age of *Name*$_j$, and s_j is the degree of similarity of *Name*$_j$ and Vera. The s_j take values in the unit interval and are bounded from below by a threshold, t. This assumption is made to enhance relevance and improve reliability. Computation of the s_j will be discussed at a later point.

Similarity-based definitions of possibility and probability are based on the knowledge of the information vector, IV, and the information dataset, ID.

Definition. The possibility, q_k, that Vera's age is a_k, Poss(Age = a_k), is defined as

$$q_k = \mathrm{Poss}(\mathrm{Age} = a_k) = \mathrm{sup}_j(s_j | a_j = a_k)$$

Definition. The probability that Vera's age is a_k, Prob(Age = a_k), is defined as

$$p_k = \mathrm{Prob}(\mathrm{Age} = a_k) = \frac{1}{s}\left(\mathrm{sum}_j\left(s_j | a_j = a_k\right)\right),$$

where s is the sum of the s_j.

Definition. The possibility distribution of Vera's age, Poss(Age), is expressed as

$$\mathrm{Poss}(\mathrm{Age}) = (q_1/a_1, \ldots, q_m/a_m)$$

Note. Poss(Age) is a fuzzy set.

Definition. The probability distribution of Vera's age, Prob(Age), is expressed as

$$\mathrm{Prob}(\mathrm{Age}) = (p_1/a_1, \ldots, p_m/a_m).$$

What is behind these very simple definitions? In the definition of possibility, q_k is the degree of similarity of the maximally-similar individual whose age is a_k. It is the maximal similarity that matters in computation of the possibility that Vera's age is a_k. More generally, what should be noted is that sup in possibility theory is the counterpart of sum in probability theory. What is the relationship between the concept of possibility as defined above and the concept of possibility in possibility theory? [3,9]. This is an issue which is in need of further exploration. In the definition of probability, p_k is the fraction of individuals whose age is a_k. This fraction is the relative $\sum$Count (relative cardinality) [10] of the fuzzy set of the *Name*$_j$ whose age is a_k. The traditional frequency-based definition of probability is a special case of the similarity-based definition. In the special case, similarity is perfect (degree of similarity = 1).

A simple example is coin tossing. I toss a coin and cover it with my hand. Did it fall heads or tails? I construct an information dataset by tossing the coin m times, resulting in an information dataset with m rows and one column, with similarity = 1. If this column contains k heads, the similarity-based probability distribution is: $\left(\frac{k}{m}/\text{heads}, \frac{m-k}{m}/\text{tails}\right)$.

More generally, the basic idea is this: You ask how old is Vera? I cannot give you a numerical answer, but based on my experience (world knowledge), I can give you the probability and possibility distributions of her age. This is the essence of the similarity-based concepts of possibility and probability distributions.

How can similarity be defined? There are many approaches. [2] One way is to define similarity of two vectors, $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in terms of similarities of their components. It should be noted that not all attributes are equally important. Assume that the importance of A_j is w_j, with the w_j taking values in the unit interval and adding up to unity. Assume that s_j is the degree of similarity of u_j and v_j.

Definition. The degree of similarity of u and v, $s(u, v)$, is defined as the convex combination of the s_j. More concretely,

$$s(u, v) = \sum_j w_j s_j.$$

It is important to note that knowing the possibility and probability distributions of Age, we can compute the possibility and probability that Vera is middle-aged. What is needed for this purpose are the definitions of the possibility and probability measures of a fuzzy event [7,9]. Let A be a fuzzy set in a space, U. Let q be a possibility distribution over U. The possibility measure of A, Poss(A), is defined as

$$\mathrm{Poss}(A) = \sup_u(q \text{ and } \mu_A),$$

where and = conjunction, u is a generic element of U, and μ_A is the membership function of A.

The probability measure of A, Prob(A), is defined as

$$\mathrm{Prob}(A) = \int_U p(u)\mu_A(u)du,$$

where p is a probability density function over U.

Consider an application of the above relation to the question: How old is Vera? Let μ_A be the membership function of middle-aged. Consider the fuzzy set $(q_1/a_1, \ldots, q_m/a_m)$. Then the probability that Vera is middle-aged may be expressed as

$$\mathrm{Prob}(\text{middle-aged}) = q_1\mu_A(a_1) + \cdots + q_m\mu_A(a_m).$$

2. Concluding remark

The similarity-based definitions of possibility and probability which are formulated in this very brief note are close in spirit to statistical procedures which are widely employed in medical diagnostics, profiling, pattern recognition, speech recognition and many other fields. The definitions are very simple and lend themselves to implementation through the use of machine-learning techniques. [5] In many realistic settings, perceptions of relevance and similarity are likely to be approximate rather than exact. In this case, similarity-based definitions of possibility and probability lead to fuzzy possibilities and fuzzy probabilities. Humans have a remarkable capability to reason and make decisions based on fuzzy possibilities and fuzzy probabilities, employing natural language for this purpose.

The concepts of similarity-based possibility and probability point to a need for a great deal of further exploration, particularly in the context of very large databases such as medical databases. Machine-learning-based algorithms for identification of question-relevant attributes and computation of similarities will have to be devised. Eventually, it may become expedient to employ linguistic variables for representation of possibility and probability distributions.

Acknowledgement

To the memory of my mentor, Herbert Robbins, and Marianne Belis.

References

[1] B.R. Bhat, Modern Probability Theory, New Age International, New Delhi, 2004.
[2] S.S. Choi, S.H. Cha, C. Tappert, A survey of binary similarity and distance measures, J. Syst., Cybernet. Inform. 8 (1) (2010) 43–48.
[3] D. Dubois, H. Prade, Possibility Theory, Plenum Press, New-York, 1988.
[4] E.T. Jaynes, Probability Theory: The Logic of Science, Cambridge University Press, New York, 2003.
[5] K. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
[6] D.W. Stroock, Probability Theory: An Analytic View, Cambridge University Press, New York, 2011.
[7] L.A. Zadeh, Probability measures of fuzzy events, J. Math. Anal. Appl. 23 (1968) 421–427.
[8] L.A. Zadeh, Similarity relations and fuzzy orderings, Inform. Sci. 3 (1971) 177–200.
[9] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets Syst. 1 (1978) 3–28.
[10] L.A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, Comp. Math. 9 (1983) 149–184.