Fuzziness and Probability

Based on texts from Jim Bezdek, Didier Dubois,
Bart Kosko, and Henri Prade

February 8, 1995

Shortly after the fuzzy set theory has been started in late 60's, there have been a number of claims that the fuzziness is nothing else but the probability in disguise. Probability and fuzziness are related but different concepts. Fuzziness is a type of deterministic uncertainty. It describes the *event class ambiguity*. Fuzziness measures the degree to which an event occurs, not whether it occurs. At issue is whether the event class can be unambiguously distinguished from its opposite. Probability arouses from the question whether or not an event occurs. Moreover, it assumes that the event class is crisply defined and that the law of non contradiction holds. That is, $A \cap A^c = \emptyset$. Kosko [6] shows that fuzziness occurs when the law of non contradiction (and equivalently the law of excluded middle, $A \cup A^c = X$) is violated. However, it seems more appropriate to investigate the *fuzzy probability* for the latter case [3], than to completely dismiss probability as a special case of fuzziness [6]. In essence, whenever there is an experiment for which we are not capable of "computing" the outcome, a probabilistic approach may be used to estimate the likelihood of a possible outcome belonging to an event class.

A fuzzy probability extends the traditional notion of a probability when there are the outcomes that belong to several event classes at the same time but to different degrees. The fuzziness and probability are orthogonal concepts that characterize different aspects of human experience. Hence, it is important to note that neither fuzziness nor probability govern the physical processes in Nature. They are *introduced* by humans to compensate for their own limitations. Detailed theoretical discussion of the relationships between the fuzziness and probability may be found in [3].

Next we look at two examples that show a difference between fuzziness and probability.

RUSSELL'S PARADOX

That the laws of non contradiction and excluded middle can be violated was pointed out by Bertrand Russell and "his barber". Russell's barber is a bewhiskered man who lives in a town and shaves a man *if and only if* he does not shave himself. The question is: who shaves the barber? If he shaves himself, then by definition he does not. But if he does not shave himself, then by definition he does. So he does and he does not. Hence a contradiction or "paradox". Gaines [4] observed that this paradoxical situation can be numerically "resolved" as follows.

Let S be the proposition that the barber shaves himself and not-S that he does not. Since S implies not-S and vice versa, the two propositions are *logically equivalent*, i.e. S=not-S. Equivalent propositions have the same truth values:

$$t(S) = t(\text{not-}S) = 1 - t(S). \tag{1}$$

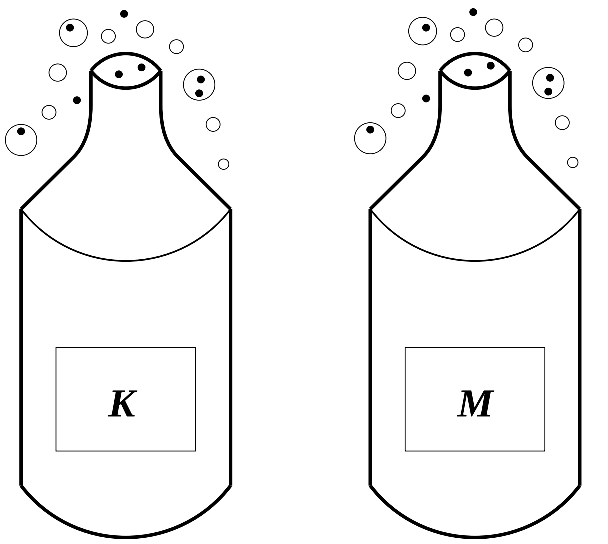
Solving for *t(S)* gives the midpoint of the truth interval [0,1]: *t(S)*=0.5. In the bivalent case this value cannot be rounded off, so the paradox occurs. However, in the fuzzy case the result is a half-truth which geometrically corresponds to the center of a fuzzy cube. The fuzzy resolution of the paradox only uses the fact that the truth values are equal. The midpoint value of 0.5 comes from the structure of the problem and the order-reversing effect of negation.

The fuzzy set theory allows for an event class to coexist with its opposite at the same time, but to different degrees, or in the case of paradox to the same degree different from zero or one. In the bivalent case, such situation is "impossible," has probability zero. The modern set theory alleviates this problem by creating sets in well defined discrete steps. However, by doing so it becomes limited to purely mathematical applications that are often far from real-world situations. Namely, the paradoxical and fuzzy situations are neglected or excluded from analysis by rounding off to the nearest vertex of a fuzzy cube. An interesting analysis of a semantic paradox using self-reference and Chaos is presented in [5].

MISLEADING SIMILARITIES

There are many similarities between fuzziness and probability. The largest, but superficial and misleading, similarity is that both systems quantify uncertainty with numbers in the unit interval [0,1]. This literarily means that both systems describe and quantify the uncertainty *numerically*. The structural similarity arousing from the lattice theory [2] is that both systems algebraically manipulate sets and propositions associatively, commutatively, and distributively. These similarities are misleading because a key distinction comes from what the two systems are trying to model.

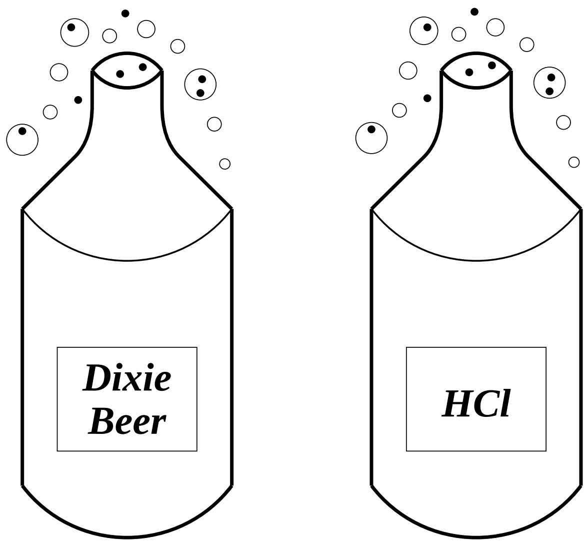
For example, [1], let L={set of all liquids}, and let fuzzy subset $\mathcal{P}$={all **potable** liquids}. Suppose you had been in the desert for a week without drink and you came upon two bottles marked K and M as in Fig. 1.



$\mu_{\mathcal{P}}(K) = 0.91$ $\quad$ $\Pr(M \in \mathcal{P}) = 0.91$

Figure 1: A pair of bottles for the weary traveler.

Confronted with this pair of bottles, and given that you must drink from the one that you chose, which would *you* choose to drink from? Most people, when presented with this experiment, immediately see that while K could contain, say, swamp water, it would not (discounting the possibility of a Machiavellian fuzzy modeler) contain liquids such as hydrochloric acid. That is, *membership* of 0.91 means that the contents of K are fairly similar to perfectly potable liquids, e.g. pure water. On the other hand, the *probability* that M is potable "equals 0.91" means that over a



$\mu_{\mathcal{P}}(K) = 0.91$ $\quad\quad$ $\Pr(M \in \mathcal{P}) = 0$

Figure 2: A pair of bottles for the weary traveler (unmasked).

long run of experiments, the contents of M are expected to be potable in about 91% of the trials. In the other 9% the contents will be deadly – about 1 chance in 10. Thus, most subjects will opt for a chance to drink swamp water.

Another distinction is in idea of *observation*. Suppose that we examine the contents of K and M and discover them to be as shown in Fig. 2. Note that, *after observation*, the membership value for K is unchanged while the probability value for M drops from 0.91 to 0.

Clearly, the two models possess different kinds of information: fuzzy memberships, which quantify similarities of objects to imprecisely defined properties; and probabilities, which provide information on expectations over a large number of experiments.

References

[1] J.C. Bezdek and P.K. Sankar, editors. *Fuzzy Models For Pattern Recognition: Methods That Search for Structures in Data*, New York, 1992. IEEE Press.

[2] G. Birkhoff. *Lattice Theory*. American Mathematical Society, New York, 1948.

[3] D. Dubois and H. Prade. Fuzzy sets and probability: Misunderstandings, bridges and gaps. In *Proc. of the Second IEEE Inter. Conf. on Fuzzy Systems*, volume 2, pages 1059–1068, San Francisco, CA, March 28-April 1 1993. IEEE.

[4] B.R. Gaines. Precise past, fuzzy future. *Intern. J. of Man-Machine Studies*, 19:117–134, 1983.

[5] P. Grim. Self-reference and chaos in fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 1(4):237–253, 1993.

[6] B. Kosko. Fuzziness vs. probability. *Int. J. General Systems*, 17:211–240, 1990.