HOW TO COMBINE PROBABILISTIC AND FUZZY UNCERTAINTIES IN FUZZY CONTROL

Extended Abstract

Hung T. Nguyen[1], **Vladik Kreinovich**[2], **Bob Lea**[3]

[1] *Department of Systems Science, Tokyo Institute of Technology*
4259 Nagatsuta-cho, Midori-ku, Yokohama 227 Japan

[2] *Computer Science Department, University of Texas at El Paso, El Paso, TX 79968*

[3] *Mail Box PT4, NASA Johnson Space Center, Houston, TX 77058*

Traditional and fuzzy control correspond to two different cases: traditional control is applicable when we have a precise knowledge of the controlled system, and possible uncertainties are described in probabilistic terms (as a *noise*). Fuzzy control (see, e.g., [S85, B91, L90]) is applicable in the situation when uncertainty is described in terms of natural language like "small", "to some extent", etc.

They both work fine. But there is a problem: suppose that we have a situation where little is known, and we have to apply fuzzy control. Since we do not know the system, the resulting control is far from being perfect. While controlling the system, in course of time we get some experience, and we can extract some statistical information from our experience. Since we now know more about the controlled system, we would like to use this additional statistical knowledge to improve the control strategy. How to do it?

At present there are no known ways to do it, and the only suggestion is to wait until we have enough information for applying traditional control theory, and then find an optimal control and switch to this control.

So, we need a method to "translate" probabilistic knowledge into fuzzy terms. In the present report, we propose such a method.

Our basic idea is as follows. In fuzzy control, we start with the uncertainty values that characterize our degree of belief that, say 0.3 is small, or that 10 is big. Where do we get these degrees of belief from? One of the possibilities is to estimate this degree as a frequency (probability), e.g., by asking several experts whether they consider 0.3 to be small or not, and then estimating the desired degree as M/N, where N is the total number of experts whom we asked, and M is the number of those who answered "yes". Probabilistic estimates are desirable, because we know how to handle probabilities. But they are not always feasible, e.g., if we are interested in the opinion of a single expert, it is difficult to get a probabilistic estimate. In such cases, to get a numerical value, we ask an expert to quantify his degree of belief, say, on a scale from 0 to 10, and then take $D/10$ as the desired estimate, where D is what an expert chose.

The readers who ever answered any sociological or psychological tests, will agree that this is the main way to quantify such vague (fuzzy) notions as "degree of satisfaction", etc.

How to translate probabilities into this kind of a scale? There are theorems [S54] that prove that optimal decision making has to deal with probabilistic estimates. There are a lot of psychological experiments that show that although people do not always consciously estimate probabilities when they make decisions, but the resulting decisions are very close to the ones that could have been obtained from the probabilities.

These two facts justify the following assumption: the scaling is based on the probabilistic estimates. From this assumption, we deduce the relationship between the position on the scale and the probability, and then we use this relationship to transform a probability into the scale position.

By saying that the scaling is based on probabilistic estimates we mean the following. When someone asks an expert to estimate his degree of belief in some statement A (e.g., that 0.3 is negligible), he recollects all the cases in which 0.3 (or a value that is close to 0.3) was tested, and figures out when 0.3 proved to be really negligible, and when the difference of 0.3 caused important changes. Suppose that totally he recalls n cases, and in m of them 0.3 was negligible. Then a reasonable estimate for the probability p (the probability that 0.3 is negligible) is $f = m/n$.

The real probability p can be different from f and, therefore, different values of m/n can correspond to one and the same probability. Indeed, if we have a sequence of n independent events with probability p, then the mathematical expectation of m/n is p, and the standard deviation σ of m/n is $\sqrt{p(1-p)/n}$. We can now apply a "$3\sigma-$ rule" from mathematical statistics, and conclude that for a given f, all the values p such that $|f-p| \leq 3\sqrt{p(1-p)/n}$ are possible. Therefore, the estimates m/n and m'/n can correspond to one and the same probability p if there exists a probability p such that $|m/n - p| \leq 3\sqrt{p(1-p)/n}$ and $|m/n - p| \leq 3\sqrt{p(1-p)/n}$. We say that the estimates are *different* if there is no such p. Now, we are ready to form a scale: we take 0 as the first element f_0 of this scale; for the second element f_1, we take the smallest estimate that is different from 0; for f_2, we take the smallest estimate that is different from both 0 and f_1, etc. Suppose that there are totally k elements on this scale. Then, when we must estimate our degree of belief on a scale from 0 to k, we recall n cases, estimate $f = m/n$, and produce k for which $f_k \leq m/n < f_{k+1}$.

In general, we do not have to recall too many cases: if we are asked for an estimate on a scale from 0 to k, we find n, for which this scale has k elements, recall n cases, and produce the resulting estimate.

Let us denote by $f(p)$ the value of this scale that corresponds to a probability p. This is not a uniform scale, because the distance between two consequent elements p and $p + \Delta p$ on this scale is proportional to $\sqrt{p(1-p)}$. In other words, $\Delta p \sim \sqrt{p(1-p)}$ leads to $\Delta f(p) = const$. For small Δp, we get $\Delta f(p) \approx f'(p)\Delta p$. Therefore, from $\Delta f(p) = const$ and $\Delta p \sim \sqrt{p(1-p)}$, we conclude that $f'(p)\sqrt{p(1-p)} = const$. This differential equation leads to $f(p) = 1/2 + 1/\pi\, arcsin(2p-1)$.

Conclusion. If it is desirable to add probabilistic knowledge to a fuzzy knowledge base, then it is necessary to transform all the probabilities p into fuzzy values $f(p) = 1/2 + 1/\pi\, arcsin(2p-1)$.

We also analyze what membership functions result from applying this procedure to standard probabilistic distributions (e.g., Gaussian), and how these functions affect the resulting fuzzy control. We also analyze what probabilistic distributions correspond to the standard membership functions (e.g., triangular or trapezoid ones).

Acknowledgements. This work was supported by a NSF Grant No. CDA-9015006, NASA Research Grant No. 9-482 and the Institute for Manufacturing and Materials Management grant. The work of the first named author was carried out at the Laboratory for International Fuzzy Engineering Research, chair of fuzzy theory, Japan.

REFERENCES

[B91] H. R. Berenji. *Fuzzy logic controllers.* In: *An Introduction to Fuzzy Logic Applications in Intelligent Systems* (R. R. Yager, L. A. Zadeh. eds.), Kluwer Academic Publ., 1991.

[L90] C. C. Lee. *Fuzzy logic in control systems: fuzzy logic controller.* IEEE Transactions on Systems, Man and Cybernetics, 1990, Vol. 20, No. 2, pp. 404–435.

[S54] L. J. Savage. *The foundations of statistics.* Wiley, New York, 1954.

[S85] M. Sugeno (editor). *Industrial applications of fuzzy control*, North Holland, Amsterdam, 1985.