LA-UR-98-1556

Approved for public release; distribution is unlimited.

Title: The Use of Fuzzy Control System Methods for Characterizing Expert Judgment Uncertainty Distributions

Author(s): Ronald E. Smith, Jane M. Booker, Thomas R. Bement, William J. Parkinson, Mary A. Meyer
Los Alamos National Laboratory

Mohammad Jamshidi
NASA Center for Autonomous Control Engineering

Submitted to: 1998 PSAM Conference

The Use of Fuzzy Control System Methods for Characterizing Expert Judgment Uncertainty Distributions

Ronald E. Smith, Jane M. Booker, Thomas R. Bement,
William J. Parkinson, Mary A. Meyer
Los Alamos National Laboratory
Los Alamos, NM, USA
Mohammad Jamshidi
NASA Center for Autonomous Control Engineering
University of New Mexico
Albuquerque, NM, USA

1 Introduction and Background

Fuzzy control system techniques are used to sythesize systems, often including their expert operators, for enhanced control of processes and systems. These techniques can be especially useful in applications involving highly nonlinear systems or systems whose mathematical models are either inaccurate or unavailable. The control system maps observed plant output parameter values into required control actions, or plant inputs. In a fuzzy control system, these observed plant outputs are transformed into degrees of membership in fuzzy plant-output sets via output membership functions. If-Then rules transform these degrees of output membership into weights associated with corresponding plant-input sets. The input sets are characterized by input membership functions. The set of possible control actions, the control-action set, is characterized by a weighted combination of the corresponding input membership functions. The precise control action is determined via a defuzzification process such as selecting the centroid of the control-action set, based on the combined input membership function that describes the control-action set [1].

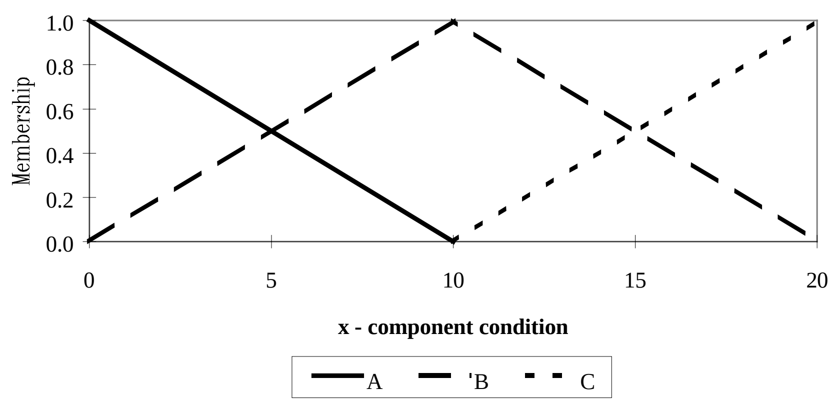
A similar process can be applied to the development of uncertainty distributions in applications such as probabilistic risk assessment, probabilistic safety assessment, and reliability analysis. For instance, the plant-output parameters used by the control system may become component condition, and the control-action may become the predicted component response or performance.

Fuzzy logic methods permit experts to assess parameters affecting performance of components/systems in natural language terms more familiar to them (e.g., "high," "good," etc.). Recognizing that there is a cost associated with obtaining more precise information, our particular interest is in cases where the relationship between the condition of the system and its performance is not well understood, especially for some sets of possible operating conditions, and where developing a better understanding is very difficult and/or expensive. The methods allow the experts to make use of the level of precision with which they understand the underlying process [2].

We consider and compare various methods of formulating the process just described, with an application in reliability analysis where expert information forms a significant (if not sole) source of data for reliability analysis. The flow of information through the fuzzy-control-systems based analysis is studied using a simple, hypothetical problem which mimics the structure used to elicit expert information in Parse (such as in NUREG 1150 [3]). We also characterize the effect of using progressively more refined information and examine the use of fuzzy-based methods as data pooling/fusion mechanisms.

2 The Fuzzy System Formulation

Consider a system with one component, which can influence performance of the system. The component is subject to wear, potentially degrading performance. For a given condition level (analogous to plant output), performance degradation will be variable, and the range of possible performance levels is analogous to the control-action set. Figure 1 shows membership functions for three hypothetical component-condition sets and three performance-level sets. The notation N(mean, standard deviation) is used for the performance-level functions which are normal distributions without the scale factor so that they range from 0 to 1.



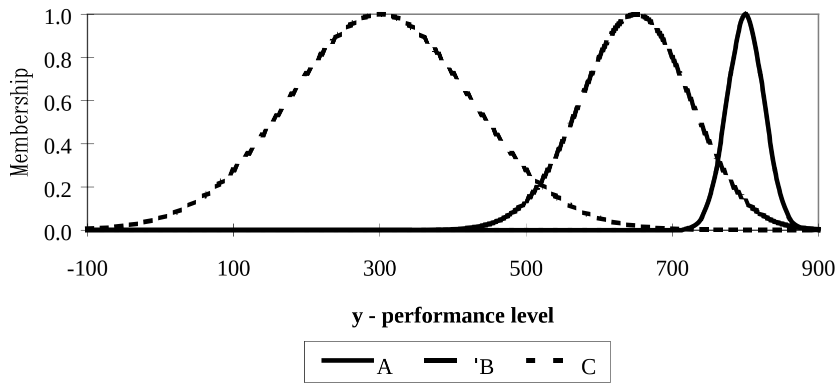


Figure 1. Component-condition and performance-level sets for 3 membership functions

Three rules define the condition/performance relationship: if condition is A, then performance is A; if condition is B, then performance is B; and if condition is C, then performance is C. If component condition is x = 4.0, then x has membership of 0.6 in A and 0.4 in B. Using the rules, the defined component component-condition membership values are mapped to performance-level weights and performance-level set A, N(800,25), characterizes the range of performance values with a weight of 0.6 and the membership function for performance-level set B, N(650,75), characterizes the range of performance values with a weight of 0.4. In fuzzy control system methods, the membership functions for performance-level sets A, N(800,25), and B, N(650,75), are combined based on the weights 0.6 and 0.4. This combined membership function can be bimodal and can be used to form the basis of an uncertainty distribution for characterizing performance for a given condition level.

Departing from standard fuzzy systems methods, we normalize the combined performance membership function so that it integrates to 1.0. The resulting function, f(y|x), is the performance, y, uncertainty distribution corresponding to the situation where component condition is equal to x. Figure 2 is the cumulative distribution function, CDF, form of the uncertainty distribution, F(y|x). If performance must exceed some threshold, T, in order for the system to operate successfully, the reliability of the system for the situation where component condition is equal to x can be expressed as R(x) = 1- F(T|x). As illustrated in Figure 2, a threshold of T = 550 corresponds to a reliability of R(4.0) = 0.925. Unless otherwise stated, the performance membership functions or combined performance membership functions will refer to normalized functions.

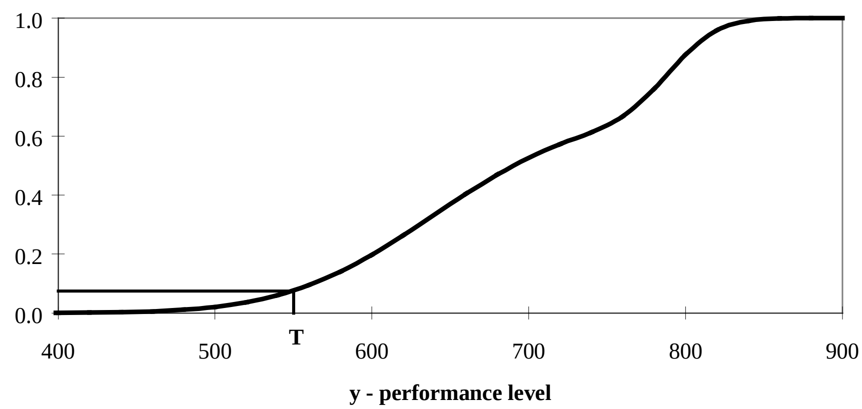


Figure 2. Performance uncertainty CDF

3 Study Description

We now consider a set of hypothetical situations to investigate how well the fuzzy characterization of uncertainty outlined above describes an unknown relationship between condition and performance. The true relationship between condition, x, and performance, y, unknown in actual applications, is specified using a series of normal distributions whose means and variances are calculated for $0 \le x \le 20$. In Figure 3, the mean (shown as the solid line) and the standard deviation of y given x are

$$\mu_y = 800 - 25x + 2x^2 - 0.1x^3, \quad \sigma_y = 25 + 5x . \tag{1}$$

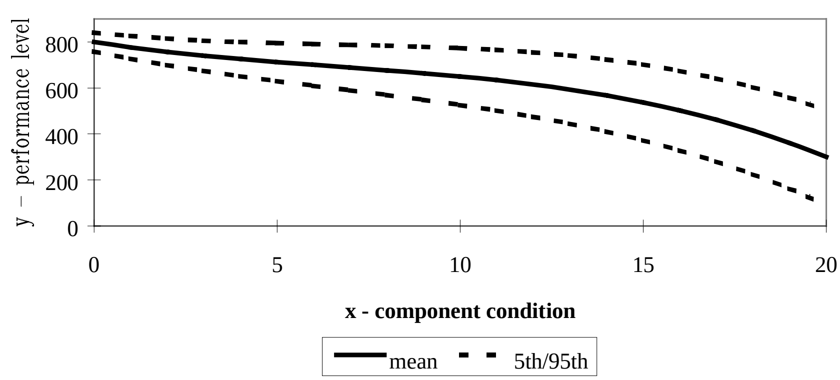


Figure 3. Underlying true relationship between condition and performance

To simplify the study, the fuzzy rules that map condition into performance are exact, meaning input membership function A maps exactly into output membership function A, a N(800,25) in this case. For the first step of the study, it is assumed that the condition/performance relationship is known exactly at three values. That is, if condition is equal to one of these values, then the performance membership function is the corresponding normal distribution specified in (1). The number of evenly spaced membership functions is then varied between three and eleven to determine how many are necessary to capture the underlying true normal distributions. From Figure 1, it is seen that for x = 0, 10, and 20, the output membership normal functions are exactly specified, with no mixtures. At all other values of x, mixtures result, as seen in Figure 2.

Formulating a mixture distribution can result in a "valley" or trough near the centroid, giving little density in the place where the central estimate is taken. A distribution from a linear combination of the two membership distributions results in a smoother transition, and a linear combination of two normals is another normal, making calculations easy. Therefore, these two combination approaches were used to calculate the uncertainty distributions for performance.

4 Results

To determine how well the resulting uncertainty distribution for performance matches the underlying true normal, goodness-of-fit tests [4] were applied for half-integer values of x on $0 \le x \le 20$. The Kolmogorov test for normality is shown to be quite powerful [5] against all alternative distributions. Figure 4 shows where the deviations from the specified underlying normals occurred for the mixture of distributions approach, using a 1% level of significance.

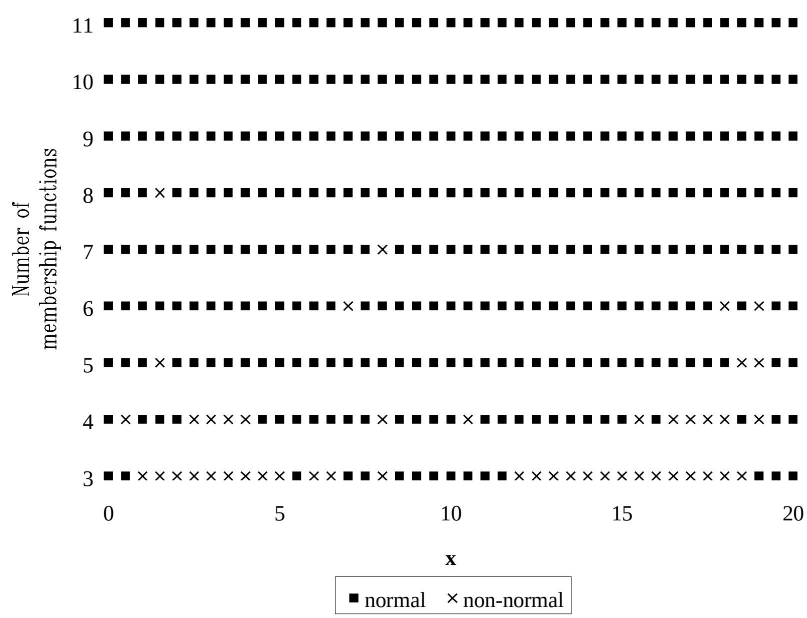


Figure 4. Deviations from normality for mixture performance distributions

Another goodness-of-fit test was developed using the Kullback-Leibler information,

$$f(x) \ln(f(x) / g(x)) \, dx \quad (2)$$

where the density of the specified normal is g(x) and the performance distribution is f(x). This test produced similar results to those in Figure 4.

Results for the linear combinations are similar to the mixture approach. In both simple cases presented here, a minimum of five membership functions is indicated for close matching to the underlying normals.

5 Conclusions and Further Studies

Although this study indicates that a minimum of five membership functions sufficiently captures the underlying true normals for either the mixture of distributions or linear combinations approaches, further studies are needed to determine the influence of other effects on this result. It is common practice to represent the expert's knowledge using triangular distributions rather than normals. The effect of using non-normal distributions for the component-condition sets and for the performance-level sets and the underlying true distributions is not known. In this study equidistant spacing of the condition membership functions were used. Common sense dictates that the experts should be able to provide better spacing of the membership functions for the component-condition sets and for the performance-level sets according to their knowledge. Such optimal spacing should result in more accurate results than indicated in this study.

In addition, work is underway to develop better goodness-of-fit tests such as those based on Kullback-Leibler information [5] for determining how well the resulting distributions match the underlying truth. Other measures of goodness-of-fit may be more useful for a given application. For example one might be interested in representing the 5^{th} and 95^{th} percentiles accurately, rather than the entire distribution.

Another area of study is multiple-input, multiple-output systems — an area where fuzzy control systems have been used successfully. One such example is a multiple component system where one is interested in both performance and safety.

The replacement of the fuzzy control system approach with a probabilistic controller approach is described in [6]. This area for further study essentially involves replacing the performance membership functions with probability density functions and formulating a mixture problem.

Acknowledgment

This work was partially supported by the US Nuclear Regulatory Commission.

References

1. Jamshidi M, Vadiee N, and Ross T, eds. Fuzzy logic and control software and hardware applications. PTR Prentice Hall, Englewood Cliffs, NJ, 1993
2. Smith R, Bement T, Parkinson W, Mortensen F, Becker S and Meyer M. The Use of fuzzy control systems techniques to develop uncertainty distributions. Proc of the Jt Stat Meetings, Anaheim, CA 1997
3. US Nuclear Regulatory Commission, Office of Regulatory Research. Severe accident risks: an assessment for five US nuclear power plants. NUREG-1150, 1989
4. Conover W. Practical Nonparametric Statistics. John Wiley & Sons, New York, 1971
5. Arizono I and Ohta H. A test for normality based on Kullback-Leibler information. The Am Stat 1989; 43: 20-22
6. Laviolette M, Seaman J Jr, Barrett J, and Woodall W. A probabilistic and statistical view of fuzzy methods. Tech 1995; 37: 249-281