54

Including Probabilistic Uncertainty in Fuzzy Logic Controller Modeling Using Dempster-Shafer Theory

Ronald R. Yager and Dimitar P. Filev

Machine Intelligence Institute

Iona College

New Rochelle, NY 10801

Technical Report #MII-1309

ABSTRACT

We discuss some basic ideas from the Dempster-Shafer theory of evidence. We describe the concept of fuzzy systems modeling used in fuzzy logic control. We use the Dempster-Shafer framework to provide a machinery for including randomness in the fuzzy systems modeling process. We show how to represent additive noise in this combined framework.

1. Introduction

Fuzzy systems model (FSM) is a technique that can be used to simplify the representation of complex nonlinear relationships. It is the basic technique used in the development of the successful fuzzy logic controllers [1-3]. Using FSM one partitions the input space into regions in which one can more simply represent the output. In FSM the partitions are determined by fuzzy subsets. The use of a fuzzy partition allows for a more sophisticated consideration of the boundaries between the regions by allowing for a weighted combination of the outputs of neighboring regions. This effectively allows us to more gradually go from one output region to the next.

In this current work we extend the applicability of FSM by suggesting a methodology for including probabilistic uncertainty in the fuzzy systems model. We particularly concentrate on the inclusion of probabilistic uncertainty in the model output. The technique we suggest for the inclusion of this uncertainty is based upon the Dempster-Shafer theory of evidence [4-8] which is closely related to the theory of random sets[9]. The Dempster-Shafer approach fits nicely into the FSM technique since both techniques use sets as their primary data structure.

We first introduce some of the basic ideas from the Dempster-Shafer theory which are required for our procedure. We next discuss the fundamentals of FSM based on the Mamdani [1] reasoning paradigm. We note that in [10] Yager described the introduction of Dempster-Shafer theory into the theory of approximate reasoning which is based on a different paradigm for fuzzy reasoning. We next show how probabilistic uncertainty in the output of a system can be included in the Mamdani type fuzzy systems model using the Dempster-Shafer paradigm. We described how various types of uncertainty can be modeled using this combined FSM/D-S paradigm. We are particularly concerned with additive noise and rule uncertainty. We next discuss a modification of the basic Mamdani formalism for FSM which leads to an analytic structure for the system output [11]. One advantage of an analytic output structure is that it simplifies the process of learning the system parameters. Finally, we look at the introduction of the Dempster-Shafer representation of probabilistic uncertainty into this analytic formalism.

2. Dempster-Shafer Theory of Evidence

In this section we introduce some ideas of the Dempster-Shafer uncertainty theory [4-8]. Assume X is a set of elements. A Dempster-Shafer belief structure m, or information granule, is a collection of non-null subsets of X, A_i, $i = 1, \ldots n$, called focal elements and a set of associated weights $m(A_i)$ such that

(1) $m(A_i) \in [0, 1]$

(2) $\sum_i m(A_i) = 1$.

It should be emphasized that no restriction exists on the collection A_i except that they don't include the null set.

One semantics that can be associated with this structure has been discussed by Yager [10, 12]. Assume we perform a random experiment which can have one of n outcomes. We shall denote the space of the experiment as Y. Let P_i be the probability of the i^{th} outcome y_i. Let V be another variable taking its value in the set X. It is the value of the variable V that is of interest to us. The value of the variable V is associated with the performance of the experiment in the space Y in the following manner. If the outcome of the experiment on the space Y is the i^{th} element, y_i, we shall say that the value of V lies in the subset A_i of X. Using this semantics we shall denote the value of the variable as

V is m,

where m is a Dempster-Shafer granule, m, with focal elements A_i and weights $m(A_i) = P_i$.

A situation which manifests the above characteristic is the following. Assume we have a "wheel of fortune" divided into three colors: red, yellow and green. The spinning of the pointer

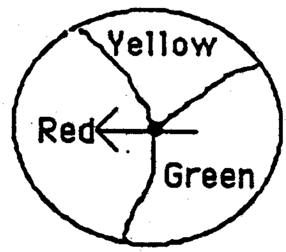

Figure #1.

results in the landing in one of the colors. The probability depends on the portion of the circle that is that color. Assume we have a vehicle whose driver is allowed to go a certain speed depending on the color in which the arrow lands. If the arrow lands in red the driver can go any speed between 10 and 30 mph, if the arrow lands in yellow the driver can go at any speed between 30 and 50 and if the arrow lands in the green the driver can go to any speed he desires (0 to 150). In this case the variable V is the speed of the car. The focal elements are A_1 = [10-30], A_2 = [30-50], A_3 = [0–150].

When our information is of the form of a Dempster-Shafer belief when attempting to try to find the probabilities associated with arbitrary subsets of X, because of the imprecision in the information, probabilities on focal elements, we can't find exact probabilities but ranges. Two measures are introduced to capture the relevant information.

Let B be a subset of X the plausibility of B, denoted Pl(B), is defined as

$$Pl(B) = \sum_{\substack{i \\ A_i \cap B \neq 0}} m(A_i)$$

The belief of B, denoted Bel(B), is defined as

$$Bel(B) = \sum_{\substack{i \\ B \subset A_i}} m(A_i)$$

It can be shown [13] that for any subset B of X

$$Bel(B) \leq Prob(B) \leq Pl(B).$$

Thus the plausibility and belief provide upper and lower bounds on the probability of a subset.

An important issue in the theory of Dempster-Shafer is the procedure for aggregating multiple belief structures on the same variable. This can be seen as a problem of information fusion. In [5] Shafer describes a procedure for combining multiple belief structures. This procedure is based upon the Dempster rule for aggregation and be seen as a kind of conjunction (intersection) of the belief structures.

Assume m_1 and m_2 are two independent belief structures on the space X, then their conjunction is another belief structure m, denoted $m = m_1 \oplus m_2$. The belief structure m is obtained in the following manner. Let m_1 have focal elements A_i, $i = 1, \dots n_1$ and let m_2 have focal elements B_j, $j = 1, \dots n_2$. The focal elements of m are all the subsets F_K of X where

(1) $F_K = A_i \cap B_j$ for some i and j

(2) $F_K \neq \Phi$.

The weights associated with each F_K is

$$m(F_K) = \frac{1}{1-T}(m_1(A_i) * m_2(B_j))$$

where

$$T = \sum_{\substack{i,j \\ A_i \cap B_j = \Phi}} m_1(A_i) * m_2(B_j)$$

Example: Assume our universe of discourse is X = {1, 2, 3, 4, 5, 6}

m_1		m_2	
$A_1 = \{1, 2, 3\}$	$m(A_1) = .5$	$B_1 = \{2, 5, 6\}$	$m(B_1) = .6$
$A_2 = \{2, 3, 6\}$	$m(A_2) = .3$	$B_2 = \{1, 4\}$	$m(B_2) = .4$
$A_3 = \{1, 2, 3, 4, 5, 6\}$	$m(A_3) = .2$		

Taking the conjunction we get

$F_1 = A_1 \cap B_1 = \{2\}$

$F_2 = A_1 \cap B_2 = \{1\}$

$F_3 = A_2 \cap B_1 = \{2, 6\}$

$F_4 = A_3 \cap B_1 = \{2, 5, 6\}$

$F_5 = A_3 \cap B_1 = \{1, 4\}$.

We also note that $A_2 \cap B_2 = \Phi$.

Since only one intersection gives us the null set then

$$T = m_1(A_2) * m(B_2) = .12$$

therefore

$$1 - T = .88.$$

Using this we calculate

$$m(F_1) = (\frac{1}{.88})(.5)(.6) = .341$$
$$m(F_2) = (\frac{1}{.88})(.5)(.4) = .227$$
$$m(F_3) = (\frac{1}{.88})(.3)(.6) = .205$$
$$m(F_4) = (\frac{1}{.88})(.2)(.6) = .136$$
$$m(F_5) = (\frac{1}{.88})(.2)(.4) = .09.$$

The above combination of belief structures can be seen to be essentially an intersection, conjunction, of the two belief structures. In [14] Yager provided for an extension of the aggregation of belief structures to any set based operation. Assume ∇ is any binary operation defined on sets, $D = A \nabla B$ where A, B and D are sets. We shall say that ∇ is an "non-null producing" operator if $A \neq \Phi$ and $B \neq \Phi \Rightarrow A \nabla B \neq \Phi$. We note that union is non-null producing but intersection is not. Assume m_1 and m_2 are two belief structures with focal elements A_i and B_j respectively. Let ∇ be any non-null producing operator. We now define the new belief structure m denoted

$$m = m_1 \nabla m_2.$$

The belief structure m has focal elements E_K where $E_K = A_i \nabla B_j$ and $m(E_K) = m_1(A_i) * m_2(B_j)$. If ∇ is **not** non-null producing we may be forced to do a process called normalization [15]. The process of normalization consists of the following

(1) Calculate

$$T = \sum_{A_i \nabla B_j = \Phi} m_1(A_i) * m(B_j)$$

(2) For all $E_K = A_i \nabla B_j \neq \Phi$ calculate

$$m(E_K) = \frac{1}{1-T} m_1(A_i) * m_2(B_j)$$

(3) For all $E_K = \Phi$ set $m(E_K) = 0$.

Example: We shall now continüe our example by considering the union of the two belief structures used in the previous example. We first note that union is a non-null producing operation. If we let

$$m = m_1 \cup m_2$$

then

	$m(E_1)$
$E_1 = A_1 \cup B_1 = \{1, 2, 3, 5, 6\}$	.3
$E_2 = A_1 \cup B_2 = \{1, 2, 3, 4\}$	.2
$E_3 = A_2 \cup B_1 = \{2, 3, 5, 6\}$	.18
$E_4 = A_2 \cup B_2 = \{2, 3, 4, 6\}$	.12
$E_5 = A_3 \cup B_1 = \{1, 2, 3, 4, 5, 6\}$	.12
$E_6 = A_3 \cup B_2 = \{1, 2, 3, 4, 5, 6\}$	.08

We can use the Dempster-Shafer structure to represent some very naturally occurring kinds of information. Assume V is a variable taking its value in the set X. Let A be a subset of X. Assume our knowledge about V is that the probability that V lies in A is "at least α." This information can be represented as the belief structure m which has two focal elements A and X and where

$$m(A) = \alpha \text{ and } m(X) = 1.$$

The information that the probability of A is exactly α can be represented as a belief structure m with focal elements A and $\overline{A}$ where $m(A) = \alpha$ and $m(\overline{A}) = 1 - \alpha$.

An ordinary probability distribution P can also be represented as a belief structure. Assume for each element $x_i \in X$ it is the case P_i is its probability. We can represent this as a belief structure where the focal elements are the individual element $A_i = \{x_i\}$ and $m(A_i) = P_i$.

Belief structures of the above kind, where the focal elements are singleton subsets, are called Bayesian [5]. For these types of structures it is the case that for any subset A of X, $Pl(A) = Bel(A)$, thus the probability is uniquely defined as a point rather than interval.

It is very natural to extend the idea of Dempster-Shafer belief structure to allow for fuzzy sets [16, 17]. We call these fuzzy Dempster-Shafer structures. We first allow the focal elements to be fuzzy subsets. In order to capture the ideas of plausibility and belief we need two ideas from the theory of possibility [18]. Assume A and B are two fuzzy subsets of X, the possibility of B given A, denoted Poss[B/A] is defined as [19]

$$\text{Poss}[B/A] = \text{Max}_i[A(x_i) \wedge B(x_i) \qquad (\wedge \text{ min})$$

The second concept is that of certainty of B given A, we denote this Cert[B/A] and define it as

$$\text{Cert}[B/A] = 1 - \text{Poss}[\overline{B}/A].$$

As shown in [16, 20] using these ideas we can extend the concepts of plausibility and belief as follows: Assume m is a belief structure on X with focal element A_i. Let B be any fuzzy subset of X. We define

$$\text{Pl}(B) = \sum_i \text{Poss}[B/A_i]m(A_i)$$

and

$$\text{Belief} = \sum_i \text{Cert}[B/A_i]m(A_i)$$

We can see that the plausibility and belief are the expected possibility and certainty of the focal elements.

The introduction of fuzzy focal elements doesn't greatly complicate the process of combination of belief structures. If ∇ is some set operation we simply use the fuzzy version of it. For example if m_1 and m_2 are belief structures and $m = m_1 \cup m_2$ then the focal element on m are

$$E_K = A_i \cup B_j$$

where

$$E_K(x) = A_i(x) \vee B_j(x) \qquad (\vee = \text{max})$$

In combining two belief structures m_1 and m_2 using some set operation ∇ we calculate the new focal elements as $E_K = A_i \nabla B_j$ and its weight as $m_1(A_i) * m_2(B_j)$. Implicit in this formulation is an assumption of independence between the belief structures. Essentially this independence is reflected in the fact that the underlying experiments generating the focal elements for each belief structure are independent. This independence manifests itself in the use of the product to calculate the new weights. That is the joint occurrence of the pair of focal elements A_i and B_j is the product of probabilities of each of them individually, $m_1(A_i)$ and $m_2(B_j)$.

In some cases we may have a different relationship between the belief structures. One very interesting case is the case we shall call synonymity. For two belief structures to be in synonymity they must essentially have their focal elements induced from the same experiment. Thus if m_1 and m_2 are two belief structures on X that are in synonymity they should have the same number of focal elements with the same weight. Thus the focal elements of m_1 are A_i, $i = 1, \ldots n$, and those of m_2 are $B_j = i = 1, \ldots n$ then $m_1(A_i) = m_2(B_i)$. If ∇ is any non-null producing set operator then if $m = m_1 \nabla m_2$ the focal elements of m are

$$E_i = A_i \nabla B_i$$

where $m(E_i) = m(A_i) = m(B_i)$.

If ∇ is not non-null producing we must use a normalization process if necessary.

3. Fuzzy Systems Modeling

Fuzzy systems modeling has shown itself to be an important tool for the development of intelligent systems, especially in the area of control.

Assume we have a complex, nonlinear multiple input single output relationship. (See Figure #2.)

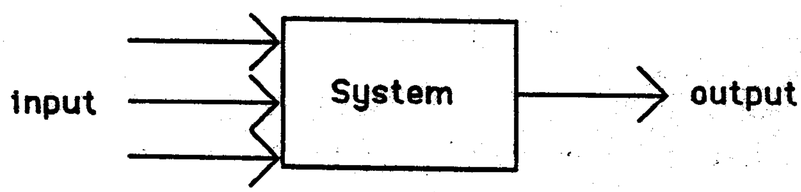

Figure #2: Basic System

The technique of fuzzy systems modeling allows us to represent the model of this system by partitioning the input space. Thus if $U_1, \ldots U_r$ are the input variables and V is the output variable we can represent the non-linear function by a collection n of "rules" of the form

When U_1 is A_{i1} and U_2 is $A_{i2}, \ldots$ and U_r is A_{ir} then V is D_i

where if X_j is the universe of discourse of U_j then A_{ij} is a fuzzy subset of X_j and with Y the universe of discourse of V then D_i is a fuzzy subset of Y.

In the preceding rule form the antecedent specifies a condition that if met allows us to infer that the possible value for the variable V lies in the consequent subset D_i. For each rule the antecedent defines a fuzzy region of the input space, $X_1 \times X_2 \times ... \times X_m$, such that if the input lies in this region the consequent holds. Taken as a collection the antecedents of all the rules form a fuzzy partition of the input space. A key advantage of this approach is that by partitioning the input space we can allow simple functions to represent the consequent.

The process of finding the output of a fuzzy systems model for a given values of the input is called the "reasoning" process. The choice of the term reasoning as opposed to solving is a result of logical structure in the which the fuzzy model is rooted. The most commonly used method for reasoning with fuzzy systems models is the Mamdani-Zadeh paradigm. [21, 22]

Assume the input to a fuzzy system model consists of the value $U_j = x_j$ for $j = 1, \ldots r$. The procedure for reasoning used in the Mamdani-Zadeh method consists of the following steps:

1. Calculate the firing level of each rule τ_i

$$\tau_i = \mathrm{Min}_j[A_{ij}(x_j)]$$

2. Calculate the output of each rule as a fuzzy subset F_i of Y where

$$F_i(y) = \mathrm{Min}[\tau_i , D_i(y)]$$

3. Aggregate the individual rule outputs to get a fuzzy subset F and Y where

$$F(y) = \mathrm{Max}_i[F_i(y)]$$

The object F is a fuzzy subset of Y indicating the fuzzy output of the system. The final step of the process is to get a singleton value for V representative of the set F. This fourth step is usually called the defuzzification process, Yager and Filev [23, 24] have investigated this issue in considerable detail. The most commonly used procedure for defuzzification process is the center of gravity [COA method]. Using this method we calculate the defuzzification value as

$$\bar{y} = \frac{\Sigma_i y_i F(y_i)}{\Sigma_i F(y_i)}$$

In [23] Yager and Filev describe this technique as essentially one in which we convert the output fuzzy subset to a probability distribution, by simple normalization, and then take the expected value over the space Y.

We can express the defuzzification operation in a simple vector notation. All vectors to follow have dimension c where c = Card(Y). Let [F] be a vector whose j^{th} component is $F(y_j)$ and let [Y] be a vector whose j^{th} component is y_j and let [I] be a vector with all ones. We can express

$$\bar{y} = \frac{<[Y], [F]>}{<[I], [F]>}$$

where <> indicates the inner product, $<A, B> = A^T B = B^T A$.

4. Probabilistic Uncertainty in the Mamdani Model

In the basic fuzzy systems model, the Mamdani-Zadeh model, the consequent of each rule consists of a fuzzy subset F_i. The use of a fuzzy subset implies a kind of uncertainty associated with the output of a rule. The kind of uncertainty is called possibilistic uncertainty and is a reflection of a lack of precision in describing the output. The use of this imprecision allows us to represent the complex nonlinear system in terms of a collection of simpler fuzzy rules. In this situation we shall consider the addition of a probabilistic component with the consequent value.

As we have indicated the consequent of the individual rules of a fuzzy system model (FSM) is a proposition of the form

$$V \text{ is } D_i.$$

The intent of this proposition if to indicate that the value of the output is constrained by (lies in) the subset D_i. The use of a set provides a robustness to this modeling technique.

We shall now add further modeling capacity to the fuzzy system modeling technique by allowing also for probabilistic uncertainty in the consequent.

A natural extension of the fuzzy systems model is to consider the consequent to be fuzzy Dempster-Shafer granules. Thus we shall now consider the output of each rule to be of the form

$$V \text{ is } m_i$$

where m_j is a belief structure with focal elements D_{ij} which are fuzzy subsets of the universe Y and associated weights $m_i(D_{ij})$. Thus a typical rule is now of the form

When U_1 is A_{i1} and U_2 is $A_{i2}, \ldots U_r$ is A_{ir} then V is m_i.

We note the antecedent portion of the rule is unchanged.

The inclusion of a belief structure to model the output of a rule is essentially saying that $m_i(D_{ij})$ is the probability that the output of the i^{th} rule lies in the set D_{ij}. So rather than being certain as to what is the output set of a rule we have some randomness in the rule. Parenthetically we note that $m_i(D_{ij}) = 1$ for some D_{ij} then we have the usual case introduced by Mamdani.

It should be carefully pointed out the use of a Dempster-Shafer granule to model the consequent of a rule brings with it two kinds of uncertainty. The first type of uncertainty is the randomness associated with determining which of the focal elements of m_i is in effect if the rule fires. This selection is essentially determined by a random experiment which uses the weights as the appropriate probability. The second type of uncertainty is related to the selection of the outcome element given the fuzzy subset, this is related to the issue of lack of specificity. This uncertainty is essentially resolved by the defuzzification procedure used to pick the crisp singleton output of the system.

Let us now formally investigate the workings of the Mamdani-Zadeh reasoning in this situation with belief structure consequents. Assume the input to the system are the values for the antecedent variables, $U_j = x_j$. The process for obtaining the firing levels of the individual based upon these inputs is exactly the same as in the previous situation. We recall for each rule the firing level, τ_i, is determined as follows

$$\tau_i = \text{Min}[A_{ij}(x_j)].$$

The output of each rule is a new belief structure $\hat{m}_i$ defined on Y denoted

$$\hat{m}_i = \tau_i \wedge m_i,$$

where $\hat{m}_i$ is a belief structure on Y. The focal elements of $\hat{m}_i$ are F_{ij} where F_{ij} is a fuzzy subset of Y defined by

$$F_{ij}(y) = \tau_i \wedge D_{ij}(y),$$

where D_{ij} is a focal element of the rule consequent. The weights associated with these new focal elements are obtained as

$$\hat{m}_i(F_{ij}) = m_i(D_{ij}).$$

The overall output of the system m is obtained in a manner analogous to that used in the Mamdani method. We obtain the overall system output m by taking a union of the individual rule outputs,

$$m = \bigcup_{i=1}^{n} \hat{m}_i.$$

In the earlier section we discussed the process of taking the union of belief structures. Thus for a collection $\mathfrak{F} = \{F_{ij(1)}, \ldots F_{ij(n)}\}$ where $F_{ij(i)}$ is some focal element of $\hat{m}_i$ we obtain a focal element of m,

$$E = \bigcup_{i} F_{ij(i)}$$

and

$$m(E) = \prod_{i=1}^{n} \hat{m}_i(F_{ij(i)})$$

Thus as a result of this third step we obtain a fuzzy Dempster-Shafer structure <u>V is m</u> as our output of the fuzzy system model. We shall assume the focal elements of m are the fuzzy subsets E_j, $(j = 1, \ldots q)$ with weights $m(E_j)$.

The next step in the procedure is to apply the defuzzification process to m to obtain the singleton output $\bar{y}$. The procedure used to obtain this defuzzified value is an extension of the originally described defuzzification procedure. For each focal elements E_j we calculate its defuzzified value $\bar{y}_j$ as follows

$$\bar{y}_j = \frac{\Sigma_i y_i E_j(y_i)}{\Sigma_i E_i(y_i)} = \frac{<[Y], [E_j]>}{<[I], [E_j]>}$$

We then take as the defuzzification value of m, $\bar{y}$,

$$\bar{y} = \Sigma_j \, \bar{y}_j \, m(E_j).$$

Thus $\bar{y}$ is essentially the expected defuzzified value of the focal elements of m.

The following simple example illustrates the technique just described.

Example: Consider a fuzzy systems model with two rules

If U is A_1 then V is m_1

If U is A_2 then V is m_2.

For simplicity we shall assume that both m_1 and m_2 are belief structures with two focal elements defined as follows:

$$m_1$$

$$D_{11} = \text{"about two"} = \left\{\frac{.6}{1}, \frac{1}{2}, \frac{.6}{3}\right\} \qquad m_1(D_{11}) = .7$$

$$D_{12} = \text{"about five"} = \left\{\frac{.5}{4}, \frac{1}{5}, \frac{.6}{6}\right\} \qquad m_1(D_{12}) = .3$$

$$m_2$$

$$D_{21} = \text{"about 10"} = \left\{\frac{.7}{9}, \frac{1}{10}, \frac{.7}{11}\right\} \qquad m_2(D_{21}) = .6$$

$$D_{22} = \text{"about 15"} = \left\{\frac{.4}{14}, \frac{1}{15}, \frac{.4}{10}\right\} \qquad m_2(D_{22}) = .4$$

We shall consider the input of the system to be x^* and assume that the membership grade of x^* in A_1 and A_2 are .8 and .5 respectively. Thus we have the firing levels of each rule

$$\tau_1 = A_1(x^*) = .8$$

$$\tau_2 = A_2(x^*) = .5.$$

Using this we can calculate the output of a belief structures of each rule

$$\hat{m}_1 = \tau_1 \wedge m_1$$

$$\hat{m}_2 = \tau_2 \wedge m_2.$$

Thus we get

$$\hat{m}_1$$

$$F_{11} = \tau_1 \wedge D_{11} = \left\{\frac{.6}{1}, \frac{.8}{2}, \frac{.6}{3}\right\} \qquad m(F_{11}) = .7$$

$$F_{12} = \tau_1 \wedge D_{12} = \left\{\frac{.5}{4}, \frac{.8}{5}, \frac{.6}{6}\right\} \qquad m(F_{12}) = .3$$

$$\hat{m}_2$$

$$F_{21} = \tau_2 \wedge D_{21} = \left\{\frac{.5}{9}, \frac{.5}{10}, \frac{.5}{11}\right\} \qquad m(F_{21}) = .6$$

$$F_{22} = \tau_2 \wedge D_{22} = \left\{\frac{.4}{14}, \frac{.5}{15}, \frac{.4}{10}\right\} \qquad m(F_{22}) = .4$$

We next obtain the union of these two belief structure,

$$m = m_1 \cup m_2.$$

The focal elements of m are obtained as follows:

$$E_1 = F_{11} \cup F_{21} \qquad m(E_1) = \hat{m}_1(F_{11}) * \hat{m}_2(F_{21})$$

$$E_2 = F_{11} \cup F_{22} \qquad m(E_2) = \hat{m}_1(F_{11}) * \hat{m}_2(F_{22})$$

$E_3 = E_{12} \cup F_{21}$ $\qquad m(E_3) = \hat{m}_1(F_{12}) * \hat{m}_2(F_{21})$

$E_4 = E_{12} \cup F_{22}$ $\qquad m(E_4) = \hat{m}_1(F_{12}) * \hat{m}_2(F_{22})$

Doing the above calculations we get

$$E_1 = \left\{\frac{.6}{1}, \frac{.8}{2}, \frac{.6}{3}, \frac{.5}{9}, \frac{.5}{10}, \frac{.5}{11}\right\} \qquad m(E_1) = .42$$
$$E_2 = \left\{\frac{.6}{1}, \frac{.8}{2}, \frac{.6}{3}, \frac{.4}{14}, \frac{.5}{15}, \frac{.4}{10}\right\} \qquad m(E_2) = .28$$
$$E_3 = \left\{\frac{.5}{4}, \frac{.8}{5}, \frac{.6}{6}, \frac{.5}{9}, \frac{.5}{10}, \frac{.5}{11}\right\} \qquad m(E_3) = .18$$
$$E_4 = \left\{\frac{.5}{4}, \frac{.8}{5}, \frac{.6}{6}, \frac{.4}{14}, \frac{.5}{15}, \frac{.4}{10}\right\} \qquad m(E_4) = .12$$

We now proceed with the defuzzification of the focal elements.

$$\text{Defuzzy}(E_1) = \bar{y}_1 = \frac{(.6)(1) + (.8)(2) + (.6)(3) + (.5)(9 + 10 + 11)}{.6 + .8 + .6 + .5 + .5 + .5} = \frac{19}{3.5} = 5.4$$
$$\text{Defuzzy}(E_2) = \bar{y}_2 = \frac{21.1}{3.3} = 6.4$$
$$\text{Defuzzy}(E_3) = \bar{y}_3 = \frac{24.6}{3.4} = 7.23$$
$$\text{Defuzzy}(E_4) = \bar{y}_4 = \frac{26.7}{3.2} = 8.34.$$

Finally combining these defuzzified values we get

$$\bar{y} = (.42) * 5.4 + (2.8) * 6.4 + (.18) * 7.23 + (.12) * 8.34$$
$$\bar{y} = 6.326$$

5. Some Classes of Uncertainty

The development of these fuzzy systems model with the addition of the Dempster-Shafer modules allows us to provide for the representation of different kinds of uncertainty associated with fuzzy modeling. We shall in the following discuss some of the possible uses of the new structure just introduced.

One important situation in which we can use the preceding model is in the case where we have a value $\alpha_i \in [0, 1]$ indicating the confidence or belief or strength of the i^{th} rule [25]. In this case we have a nominal rule of the form

If U is A_i then V is B_i

<u>with belief/confidence "at least α_i"</u>. Without loss of generality we have assumed single antecedents.

Using the framework developed above we can transform this rule, along with its associated confidence level into a Dempster-Shafer structure

If U is A_i then V is m_i.

In this new structure m_i is a belief structure with two focal elements, A_i and Y. We recall Y is the whole output space. The associated weights are $m_i(A_i) = \alpha_i$ and $m(Y) = 1 - \alpha_i$.

We see that if $\alpha_i = 1$ then we get the original rule while if $\alpha_i = 0$ we get a rule of the form

If U is A_i then V is Y.

Another important application is the modeling of additive noise to the systems output. Assume we have a system which has additive noise, that is the system output is $V + N_i$, where N_i is some noise.

We consider first using fuzzy systems modeling to provide output without considering the noise component. Thus we use the typical rules of the form

When U is A_i then V is B_i.

Again we call B_i the nominal output value.

We now assume that the output is contaminated by output noise N_i, thus we get

When U is A_i then V is $B_i + N_i$.

The noise component is a random component. The information about the noise resides in a possibility density function, f(x), (see Figure #3).We shall now investigate how we can effectively include this noise in our model in the same spirit as our fuzzy model.

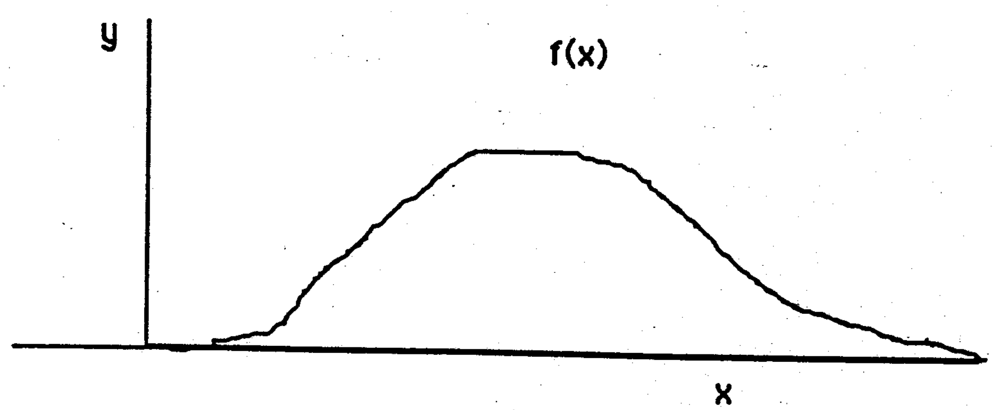

Figure #3: Possibility Density Function

We recall that the area under the curve is equal to one. Furthermore, since

$$P(x_1 \leq N \leq x_2) = \int_{x_1}^{x_2} f(x)\, dx$$

then the area under the curve between x_1 and x_2 is the probability that N lies between x_1 and x_2. (See Figure #4.)

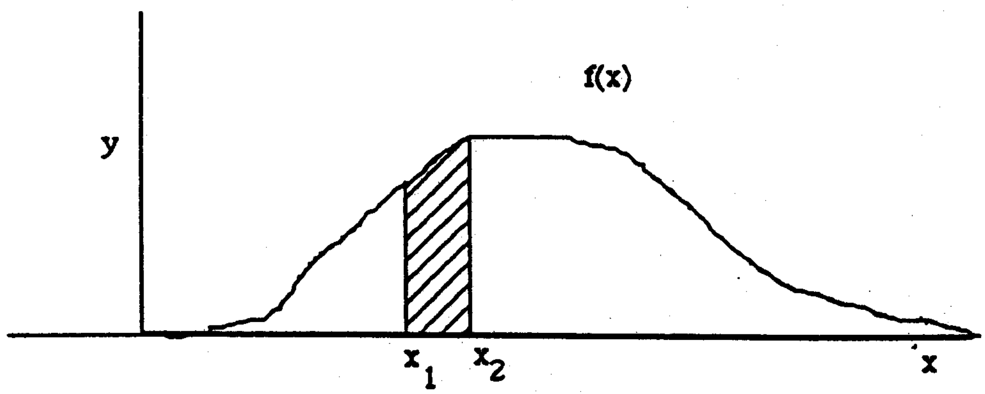

Figure #4.

One can suggest an approximation to this probability density function by a fuzzy probability distribution P. A fuzzy probability distribution P consists of a collection of fuzzy sets of real line, $Q_1, \ldots Q_m$ and associated probability distributions $P_1, \ldots P_m$, where

1) $\sum_i P_i = 1$
2) $P_i \in [0, 1]$.

Thus P_i is the probability fuzzy subset Q_i will occur. Figure #5 shows the relationship between the additive noise probability density and the fuzzy probability distribution.

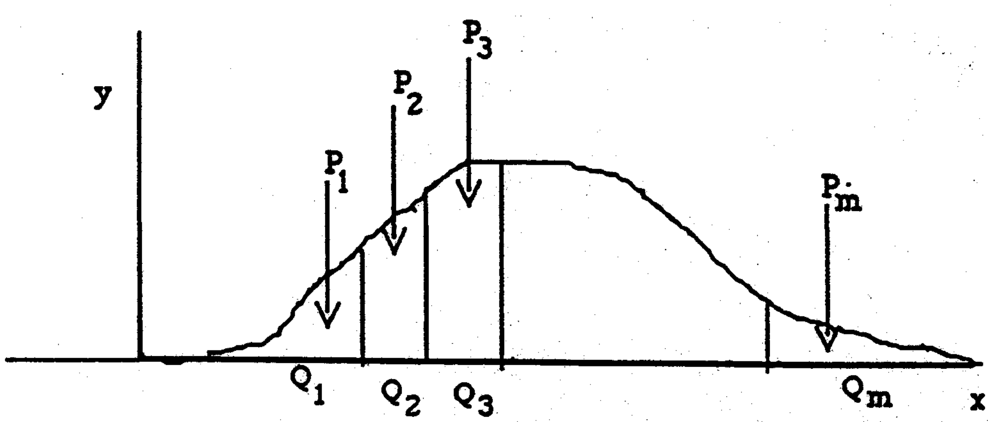

Figure #5.

In particular P_1 is the area under the curve indicating the probability that N lies in the fuzzy interval Q_1. Essentially we are formally saying

$$P_i = P(N \in Q_i) = \int_{Q_i} f(x)\, dx.$$

However our fuzzy probability distribution is only an approximation to this.

It is essentially a fuzzy partitioning of the probability density function. We now can use this fuzzy probability distribution to model the random noise N in our fuzzy systems model.

Starting with our basic rule

If U is A_i then V is $B_i + N$

we transform this into collection of rules

If U is A_i then V is $B_i + Q_1$	P_1
If U is A_i then V is $B_i + Q_2$	P_2
If U is A_i then V is $B_i + Q_m$	P_m

However it can be easily seen that this can be represented as a Dempster-Shafer belief structure.

If U is A_i then V is m_i

where the focal elements of m_i are

$F_{i1} = B_i + Q_1$	$m(F_{i1}) = P_1$
$F_{i2} = B_i + Q_2$	$m(F_{i2}) = P_2$
$F_{im} = B_i + Q_m$	$m(F_{im}) = P_m$

The operation of addition of these sets can easily be accomplished by fuzzy arithmetic [26]. We recall that if G and H are two fuzzy subsets of the real line then if

$$E = G + H$$

we have that for any $z \in$ Reals

$$E(z) = \underset{\substack{x,\, y \\ z=x+y}}{\text{Max}} [G(x) \wedge H(x)].$$

Thus the inclusion of additive noise in our model accomplished by using belief structures as the consequents of the rules. These belief structures have focal elements obtained by starting with the nominal output and then perturbing by adding the noise component.

6. Analytic Fuzzy Modeling

With the wide applications of the fuzzy systems modeling in the fuzzy logic controllers considerable interest has been focussed on the development of a more analytic and tractable representation of the reasoning process than that provided by the original Mamdani-Zadeh paradigm. This interest has been motivated to a large degree by a desire to automate the process of learning the rules from data which is difficult using max's and min's. A second motivating factor for this desire for simplification is to speed up the on-line reasoning process.

In [27-30] Yager and Filev have investigated an alternative model for the representation and reasoning process in the fuzzy system modeling paradigm. In particular they suggest the following modification to the original Mamdani-Zadeh paradigm:

I. Replace the min operation used in the process of determining the individual rule output from its firing level and consequent to using the product,

$$F_i(y) = \tau_i * D_i(y)$$

II. Replace the max operation used in the process of aggregating the rule outputs to use of an average

$$F(y) = \frac{1}{n} \sum_i F_i(y).$$

The introduction of these modifications as shown in [27-30] leads to a very simplified analytic expression for the defuzzified output, $\bar{y}$, of the fuzzy system model,

$$\bar{y} = \frac{\Sigma_i \tau_i S_i \bar{y}_i}{\Sigma_i \tau_i S_i}$$

In the above $\bar{y}$ is the overall output, $\bar{y}_i$ is the defuzzified, using the COA method, value of the i^{th} rule consequent D_i. τ_i is the firing level of the i^{th} rule. S_i is the power of the consequent fuzzy subset D_i, $S_i = \sum_j D_i(y_j)$. For continuous consequents S_i is the area under the membership function

$$S_i = \int_y D_i(y) dy.$$

If it is assumed that all the S_i are equal then

$$\bar{y} = \frac{\Sigma_i \tau_i \bar{y}_i}{\Sigma \tau_i}$$

This form of output which is close to the Takagi-Sugeno [31] structure and is in the spirit suggested by Kosko [32] has been used in many applications of fuzzy logic controllers.

In the following section we shall investigate the structure of the models we obtained using the modification described in I and II to the case where the consequent is a belief structure. In particular we shall look at the case where we have additive noise. For simplicity we shall assume that the additive noise components are point sets. Thus in this case we have a collection of r rules of the type

$$\text{If U is } A_i \text{ then V is } m_i$$

where m_i is the belief structure with focal elements D_{ij} with weights $m_i(D_{ij})$. Furthermore we assume that

$$D_{ij} = D_i \oplus u_j.$$

In the above D_i is the nominal fuzzy output of the i^{th} rule and u_j is a noise component which is assumed a crisp point thus

$$D_{ij}(z) = D_i(z - u_j).$$

We shall denote $m_i(D_i) = p_i$.

Let us use τ_i to indicate the firing level of the i^{th} rule. The output of the system is the fuzzy Dempster-Shafer granule

$$\text{V is m}$$

where m has focal elements E_q. We note that E_q is constructed by selecting one focal element from each m_i, thus we let D_{ii_q} be the focal element from m_i involved in the construction of E_q. In particular the form of E_q is

$$E_q(y) = \frac{1}{r}\left[\sum_{i=1}^{r} \tau_i D_{ii_q}(y)\right]$$

and

$$m(E_q) = \prod_{i=1}^{r} P_{ii_q},$$

where $P_{ii_q} = m_i(D_{ii_q})$.

Furthermore we can calculate the defuzzified value of E_q. In particular

$$\text{Defuzz}(E_q(y)) = \bar{y}_q = \frac{[E_q(y)]\,[Y]}{[E_q(y)]\,[I]}$$

$$\bar{y}_q = \frac{\Sigma_i\, \tau_i\, S_i\, (\bar{d}_i + u_{i(q)})}{\Sigma_i\, \tau_i\, S_i}$$

In the above S_i is the power of the set D_i, $S_i = \Sigma_j D_i(y_j)$, $\bar{d}_i$ is the defuzzified value associated with D_i and $u_{i(q)}$ is the noise component associated with the i^{th} rule in the q focal element.

The final output of this system is

$$\bar{y} = \sum_q \bar{y}_q\, m(E_q).$$

It can be seen that D_{ij} appears in a focal element with every combination of all the elements from the other rules. In particular this leaves a very simple form for the overall defuzzified value

$$\bar{y} = \frac{\Sigma_i\, \tau_i\, S_i\, (\bar{d}_i + \overline{U}_i)}{\Sigma_i\, \tau_i\, S_i}$$

where

$$\overline{U}_i = \Sigma_j\, U_{ij}\, P_{ij},$$

the expected value of the noise in the i^{th} rule.

The above formulation for $\bar{y}$ provides a nice closed analytic form for the calculation of the output of this fuzzy systems model.

We now consider the case where each rule has a measure α_i of certainty associated with it. Assume our knowledge base consists of a set of n rules of the form

If U is A_i then V is m_i

where m_i is a belief structure with two focal elements, D_i and Y. The weights associated with these are $m_i(D_i) = \alpha_i$ and $m_i(Y) = 1 - \alpha$. We recall that Y is the whole output space. In the previous section we suggested that this corresponds to a situation in which we have α_i support associated with the i^{th} rule. That is, α_i is the belief we have in rule i.

We shall look at the form of the output of such a fuzzy system model using the simplified reasoning paradigm.

In the most general case because of the S_i term the formulation of such an analytic model is difficult. However, if we consider the special situation where the sets D_i and Y are represented by point values, $\overline{y}_i$ and y_o respectively we get a nice simplified closed form for the system output

$$\overline{y} = \frac{\Sigma_i (\tau_i)((\alpha_i \overline{y}_i) + y_o(1 - \alpha_i))}{\Sigma_i \tau_i}$$

$$\overline{y} = \frac{y_o \Sigma\tau_i + \Sigma\tau_i\alpha_i(\overline{y}_i - y_o)}{\Sigma_i \tau_i}$$

7. Conclusion

We have suggested a methodology which can be used for the inclusion of randomness in the framework of fuzzy systems modeling. We used the Dempster-Shafer theory of evidence to accomplish this task.

8. References

[1]. Mamdani, E. H. and Assilian, S., "An experiment in linguistic synthesis with a fuzzy logic controller," Int. J. of Man-Machine Studies 7, 1-13, 1975.

[2]. Sugeno, M., Industrial Applications of Fuzzy Control, North-Holland: Amsterdam, 1985.

[3]. Tong, R. M., "An annotated bibliography of fuzzy control," in Industrial Applications of Fuzzy Control, Sugeno, M. (Ed.), North-Holland: Amsterdam, 249-269, 1985.

[4]. Dempster, A. P., "A generalization of Bayesian inference," Journal of the Royal Statistical Society, 205-247, 1968.

[5]. Shafer, G., A Mathematical Theory of Evidence, Princeton University Press: Princeton, N.J., 1976.

[6]. Shafer, G., "Belief functions and possibility measures," in Analysis of Fuzzy Information, Vol 1: Mathematics and Logic, edited by Bezdek, J. C., CRC Press: Boca Raton: Florida, 1987.

[7]. Smets, P., "Belief Functions," in Non-Standard Logics for Automated Reasoning, Smets, P., Mamdani, E.H., Dubois, D. & Prade, H. (eds.), London: Academic Press, 253-277, 1988.

[8]. Yager, R. R., Fedrizzi, M. and Kacprzyk, J., Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons: New York, (To Appear).

[9]. Goodman, I. R. and Nguyen, H. T., Uncertainty Models for Knowledge-Based Systems, North-Holland: Amsterdam, 1985.

[10]. Yager, R. R., "Reasoning with uncertainty for expert systems," Proc. of the Ninth International Joint Conference on Artificial Intelligence, Los Angeles, 1295-1297, 1985.

[11]. Yager, R. R. and Filev, D. P., "Unified structure and parameter identification of fuzzy models," IEEE Transactions on Systems, Man and Cybernetics, (To Appear).

[12]. Yager, R. R., "On implementing usual values in Uncertainty," in Artificial Intelligence 2, Lemmer, J.F. & Kanal, L.N., (Eds.), New York: Elsevier Science Publishers, 209-217, 1988.

[13]. Dempster, A. P., "Upper and lower probabilities induced by a multi-valued mapping," Ann. of Mathematical Statistics 38, 325-339, 1967.

[14]. Yager, R. R., "Arithmetic and other operations on Dempster-Shafer structures," Int. J. of Man-Machine Studies 25, 357-366, 1986.

[15]. Yager, R. R., "Quasi-associative operations in the combination of evidence," Kybernetes 16, 37-41, 1987.

[16]. Yager, R. R., "Entropy and specificity in a mathematical theory of evidence," Int. J. of General Systems 9, 249-260, 1983.

[17]. Yen, J., "Generalizing the Dempster-Shafer theory to fuzzy sets," IEEE Transactions on Systems, Man and Cybernetics 20, 559-570, 1990.

[18]. Zadeh, L. A., "Fuzzy sets and information granularity," in Advances in Fuzzy Set Theory and Applications, Gupta, M.M., Ragade, R.K. & Yager, R.R. (eds.), Amsterdam: North-Holland, 3-18, 1979.

[19]. Zadeh, L. A., "Fuzzy sets as a basis for a theory of possibility," Fuzzy Sets and Systems 1, 3-28, 1978.

[20]. Yager, R. R., "On the Dempster-Shafer framework and new combination rules," Information Sciences 41, 93-137, 1987.

[21]. Lee, C. C., "Fuzzy logic in control systems: fuzzy logic controller, Part 1," IEEE Trans. on Systems, Man and Cybernetics 20, 404-418, 1990.

[22]. Lee, C. C., "Fuzzy logic in control systems: fuzzy logic controller, Part II," IEEE Trans. on Systems, Man and Cybernetics 20, 419-435, 1990.

[23]. Yager, R. R. and Filev, D. P., "On the issue of defuzzification and selection based on a fuzzy set," Fuzzy Sets and Systems, (To Appear).

[24]. Filev, D. and Yager, R. R., "A generalized defuzzification method under BAD distributions," International Journal of Intelligent Systems 6, 687-697, 1991.

[25]. Yager, R. R., "On considerations of credibility of evidence," Int. J. of Approximate Reasoning 7, 45-72, 1992.

[26]. Dubois, D. and Prade, H., "Fuzzy numbers: An overview," in Analysis of Fuzzy Information Vol 1: Mathematics and Logic, Bezdek, J.C. (ed.), Crc Press: Boca Raton, Fl, 3 - 39, 1987.

[27]. Yager, R. R. and Filev, D. P., "Generalizing the modeling of fuzzy logic controllers by parameterized aggregation operators," Technical Report #MII-1219, Machine Intelligence Institute, Iona College, 1992.

[28]. Yager, R. R. and Filev, D. P., "Adaptive defuzzification for fuzzy logic controllers," BUSEFAL 49, 50-58, 1992.

[29]. Yager, R. R. and Filev, D. P., "Fuzzy logic controllers with flexible structures," Proceedings Second International Conference on Fuzzy Sets and Neural Networks, Iizuka, Japan, 317-320, 1992.

[30]. Yager, R. R. and Filev, D. P., "Analysis of flexible structured fuzzy logic controllers," Technical Report #MII-1232 Machine Intelligence Institute, Iona College, 1992.

[31]. Takagi, T. and Sugeno, M., "Derivation of fuzzy control rules from human operators actions," Proceedings of the IFAC Symposium on Fuzzy Information, Marseille, 55-60, 1983.

[32]. Kosko, B., Neural Networks and Fuzzy Systems, Prentice Hall: Englewood Cliffs, NJ,