Engineering Applications of Artificial Intelligence 72 (2018) 93–98

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Fuzzy rule bases with generalized belief structure inputs

Ronald R. Yager

Machine Intelligence Institute, Iona College, New Rochelle, NY 10801, United States

ARTICLE INFO

Keywords:
Fuzzy systems modeling
Measure
Uncertainty
Interval value
Dempster–Shafer belief structure

ABSTRACT

We first describe the basics of fuzzy systems modeling. Fundamental to this is a collection of rules, a rule base, in which the rule antecedents are fuzzy subsets. We first look at issue of the determination of the firing level of a rule for fuzzy set inputs and the subsequent rule base output. We next consider the situation where the system input is uncertain and modeled by a Dempster–Shafer belief structure. Here our input is a collection of fuzzy subsets and the true input fuzzy set is selected based on a probability distribution over these potential input fuzzy sets. We next consider the situation where our input is modeled via a generalized belief structure where the determination of applicable input fuzzy set is modeled via a measure over these potential input fuzzy sets.

1. Introduction

Fuzzy systems modeling (Mendel, 2017; Pedrycz and Gomide, 2007) is clearly the most successful application of Zadeh's fuzzy set theory (Zadeh, 1965). Using this idea we are able to model complex input/output functions by imprecisely partitioning the input space using fuzzy sets, and then associating with each fuzzy set in this partitioning an appropriate output. One advantage of this approach is that it allows humans to more easily express the component relationships. In fuzzy systems modeling each of these input/output pairs have come to be called a rule whose antecedent is a fuzzy set and the collection of these component rules is called a rule base. The determination of the output of a fuzzy systems model for given value for the input is implemented via a very natural process. The first step is to obtain the relevance of a rule to the given input, this is called the rule firing level. The systems output is then obtained as a kind of weighted average of individual rule outputs where the weight associated with a given rule is based on its firing level for the system input. The original formalism for fuzzy systems modeling was due to Mamdani and Assilian (1975) and Mamdani (1976) however the prevalent approach to fuzzy systems modeling is based on the ideas of Takagi and Sugeno (1985). Most work on fuzzy systems modeling has focused on two types of systems input values; precise values and imprecise values captured using fuzzy sets. When the input to the system is a precise value a rule firing level is simply the membership grade of the input in the rule antecedent fuzzy sets. When the input is a fuzzy set obtaining the rule firing level involves the determination of the satisfaction of one imprecise object, fuzzy set, by another imprecise object, fuzzy set. As we shall see a reasonable value for the firing level in this case is an interval. In this work we go beyond these two situations. First we consider the situation were the system input value is modeled by a Dempster–Shafer belief structure (Dempster, 1967; Shafer, 1976; Dempster, 2008; Yager and Liu, 2008). This type of input has probabilistic uncertainty as well imprecision. Next we consider the situation were the input is modeled by generalized belief structure (Yager, 2017, 2018) here we have measure guided uncertainty (Yager, 2016) as well imprecision. In these situations where the systems input manifests uncertainty as well as imprecision the determination of rule firing level as well the systems outcome becomes more complex. We see that the novelty and benefit of this paper is that it provides tools for working with fuzzy systems models for various types of uncertain inputs.

The structure of the paper is as follows we first discuss the idea of fuzzy systems modeling and look at the determination of the firing level of a rule for fuzzy set inputs. We then consider the situation when input is uncertain and modeled by a Dempster–Shafer belief structure with fuzzy focal elements. We next look the case where our input is modeled via a generalized belief structure.

2. Basics of fuzzy systems modeling

We now describe the basic framework of fuzzy systems models (Mendel, 2017; Pedrycz and Gomide, 2007; Ross, 2010). Assume V_j for $j = 1$ to r are a collection of variables taking their values in the spaces X_j respectively, these are called the input variables. We let U be another variable taking its value in the space Y, this is called the output variable. We let W for $k = 1$ to t be another collection of variables which can contain some of the V_j. Central to the fuzzy system modeling technique is a rule base consisting of a collection of $i = 1$ to n rules of the form.

If V_1 is A_{i1} and V_2 is A_{i2} and … and V_r is A_{ir} then U is $f_i(W_1, \dots, W_t)$.

E-mail address: yager@panix.com.

https://doi.org/10.1016/j.engappai.2018.03.005
Received 31 October 2017; Received in revised form 12 February 2018; Accepted 5 March 2018
Available online 8 May 2018

Here A_{ij} is a normal fuzzy subset of the universe X_j and f_i is a function that maps into the space Y. We note that more sophisticated antecedent formulations then the simple "anding" of the input variables is possible using various aggregation operators, however for our purpose we shall use the basic "anding".

Another simplifying assumption we shall make here is with respect to the form of $f_i(W_1, \ldots, W_t)$, here we shall assume that the $f_i(W_1, \ldots, W_t)$ are constants, $f_i(W_1, \ldots, W_t) = b_i$. This formulation is assumed in most applications of the Takagi–Sugeno methodology (Takagi and Sugeno, 1985; Sugeno and Takagi, 1983). So the rule base we will work with is a collection of n rules, R_i, for $i = 1$ to n, of the form.

If V_1 is A_{i1} and V_2 is A_{i2} and … ..and V_r is A_{ir} then U is b_i.

Given information about the input variables, which we denote as V_j is Inf(j), we proceed to implement the fuzzy systems model as follows.

(1) For each antecedent component V_j is A_{ij} we calculate its satisfaction, $\tau_{ij} \in [0,1]$, by the value of V_i, Inf(j). We indicate $\tau_{ij} = \text{Val}(A_{ij}/\text{Inf}(j))$.

(2) For each rule R_i we calculate its firing level, $\tau_i = \prod_{j=1}^{r} \tau_{ij}$.

(3) Using this we obtain the system output $U = b$, where $b = \frac{\sum_{i=1}^{n} \tau_i b_i}{\sum_{i=1}^{n} \tau_i}$.

Here we are taking a weighted average of the rule outputs based on their current firing levels.

A number of variations of this approach are possible, for example for we can calculate b such that $b = b_k$ where $\tau_k = \text{Max}_i[\tau_i]$, here b is the output of strongest fired rule. This approach is particularly useful in the cases where the output space Y does not allow arithmetic operations. Another approach to the selection of b is the Random Experiment Decision, RED approach. Here we associate with each b_i a profitability $p_i = \frac{\tau_i}{\Sigma_i \tau_i}$ and select the value b by the performance of a random experiment where p_i is the probability of selecting b_i. This also works in situations in which the space Y does not allow arithmetic operations. This RED approach is particularly useful in competitive decision-making situations when we do not want a given input to always result in the same action.

In the case where we desire a more complex antecedent relationship we have available many aggregation operators (Beliakov et al., 2007) to calculate τ_i as $\tau_i = \text{Agg}_i(\tau_{i1}, \ldots, \tau_{ir})$.

3. Determination of antecedent firing levels

We now consider the determination of the firing levels for the individual antecedent components, the V_j is A_{ij}, given the knowledge Inf(j) about the value of the variable V_j. In the following, in order to avoid unnecessary notational complexity we shall, when it does not involve a loss of generality, consider one generic rule

If V_1 is A_1 and V_2 is $A_2, \ldots$ and V_n is A_n then U is b.

In this case where we know the exact value of V_j, $V_j = a_j$ then $\tau_j = A_j(a_j)$, the membership grade of a_j in A_j. Once we introduce some uncertainty and/or imprecision in the information about the value of V_j the situation becomes complex. In Yager (2015) we considered the case when our knowledge about V_j is expressed by a normal fuzzy subset E_j of X_j. Here at least one element of the space X_j has membership of one in E_j.

In Yager (2015) we provided a general formulation for the function used to model τ_j, the satisfaction of the antecedent condition V_j is A_j given V_j is E_j. We shall express this as the validity of A_j given E_j and denote τ_j as Val$(A_j/E_j]$. In the following for notational simplicity we shall suppress the j and simply use the notation Val(A/E). Thus here we are interest in the satisfaction of the condition V is A given the knowledge that V is E.

Definition. If A and E are two normal subset of X we shall say a validity operator, Val(A/E) is a credible quantification of the degree of satisfaction of V is A given V is E if Val(A/E) has the following properties:

(1) If E is a singleton, $E = \{x\}$, then Val$(A/E) = A(x)$

(2) If A is the whose space, $A = X$, then for any E, Val$(A/E) = 1$

(3) If $A \cap E = \varnothing$ then Val$(A/E) = 0$

(4) Val(A/E) is monotonic in A, if $A_1 \subseteq A_2$ then Val$(A_1/E) \leq$ Val(A_2/E).

In Yager (2015) a number of formulations for Val(A/E) were discussed. One formulation is the measure of possibility (Zadeh, 1979) here

$$\tau = \text{Val}(A/E) = \text{Poss}(A/E) = \text{Max}_x[A(x) \wedge E(x)].$$

The use of possibility has one very questionable feature if $E = X$, then Val$(A/E) = 1$. Thus we get complete satisfaction to the requirement V is A in the case when we know noting about the current value of V other than V lies in X. More generally since Poss(A/E) increases as the fuzzy subset E increases, we see that the firing level of V is A increases as the imprecision in the knowledge about V increases.

Another approach is to use the measure of certainty (Zadeh, 1979), here

$$\tau = \text{Val}(A/E) = \text{Cert}(A/E) = 1 - \text{Poss}[\bar{A}/E] = 1 - \text{Max}_x[\bar{A}(x) \wedge E(x)].$$

Here we are obtaining the firing level as the negation of possibility of not A being satisfied. Here if $E = X$, the know nothing we get $\tau = \text{Val}(A/E) = 0$.

It can be shown that Cert$(A/E) \leq$ Poss(A/E). So possibility always gives us at least as large a firing level as certainty. Here we see a reasonable conservative observation is that Cert$(A/E) \leq \tau \leq$ Poss(A/E), that is Val$(A/E) \in [\text{Cert}(A/E), \text{Poss}(A/E)]$.

One can obtain some other less conservative formulations for Val(A/E) as particular points in this interval. One notable example of Val(A/E) described in Yager (2015) is based on the proposition of elements in E that are also in A

$$\text{Prop}(A/E) = \frac{\Sigma_j A(x_j) E(x_j)}{\Sigma_j E(x_j)}.$$

In the case where $E = X$, Prop$(A/E) = \frac{\Sigma_j A(x_j)}{\Sigma_j x(x_j)} = \frac{\Sigma_j A(x_j)}{n}$ = Average membership grade of A.

In Yager (2015) the author introduced another formulation for Val(A/E) based on the Choquet integral and denoted CP(A/E). In the following we let ρ be an index function so $\rho(k)$ is the index of the element in domain X of V with the kth largest value for $A(x_i)$, here $A(x_{\rho(k)})$ has the kth largest membership grade in A. Using this ρ we define

$$\text{CP}(A/E) = \sum_{j=1}^{n} (\text{Max}_{k \leq j}[E(x_{\rho(k)})] - \text{Max}_{k \leq j-1}[E(x_{\rho(k)})]) A(x_{\rho(j)}).$$

By convention in the above for $j = 1$ we define $\text{Max}_{k \leq j-1}[E(x_{\rho(k)})] = 0$ that is $\text{MAX}_{k \leq 0}[E(x_{\rho(k)})] = 0$.

The validity operator is an interesting operator, in Yager (2015) a number of properties of this validity operator were described. If $w_1, \ldots, w_q \in [0,1]$ and $\Sigma_i w_i = 1$ then if Val$_i(A/E)$ are a collection of validity operators then the operator Val$(A/E) = \Sigma_i w_i \text{Val}_i(A/E)$ is also a validity operator. Thus the linear combination of validity operators is a valid operator. One notable example of linear combination is Val(A/E) $= \alpha$ Poss$(A/E) + \bar{\alpha}$ Cert(A/E).

If Val$_1, \ldots,$ Val$_q$ are validity operators then Val$(A/E) = \text{Max}_i[\text{Val}_i(A/E)]$ is a validity operator also $V(A/E) = \text{Min}_i[\text{Val}_i(A/E)]$ is a validity operator.

More generally if Val$_1, \ldots,$ Val$_q$ are validity operators and H is any mean aggregation operator (Beliakov et al., 2007) then Val$(A/E) = H(\text{Val}_1(A/E), \text{Val}_2(A/E), \ldots, \text{Val}_q(A/E))$ is a validity operator.