Constructing Hierarchical Fuzzy Rule Bases for Classification

Tamás D. Gedeon
School of Information Technology
Murdoch University
South St, Murdoch, Western Australia 6150

Kok Wai Wong
School of Information Technology, Murdoch University

Domonkos Tikk
Department of Telecom. & Telematics, Budapest University of Technology and Economics

Abstract:
Fuzzy rule based systems have been very popular in many control applications. However, when fuzzy control systems are used in real problems, many rules may be required. The number of rules required depends on the number of inputs and the number of fuzzy linguistic terms used. This exponential explosion of fuzzy rules can take too much computing time to solve any but the simplest problems. This paper proposes a hierarchical fuzzy system that partitions a problem for more efficient computation. The hierarchical fuzzy rule base algorithm constructs rules from data for the purpose of performing fuzzy classification. Illustration examples are also generated and the results show that this hierarchical fuzzy system can be successfully used for classification applications.

I. Introduction

In general, fuzzy control systems are still the most important applications of fuzzy theory. This is a generalised form of expert control using fuzzy sets with fuzzy rules in modelling a system. In classical fuzzy approaches from Zadeh [1] and Mamdani [2], the basic idea is to calculate the conclusion by evaluating the degree of matches from the observation that triggered one or several rules in the model.

In most fuzzy modelling or fuzzy control systems, experiments and simulations are set up to generate a set of data that best describe all the possible outcomes. After this, a human expert will create the set of fuzzy rules that best perform control or modelling. Normally, fuzzy rules generated in this manner will cover the whole universe of discourse by taking all the possibilities into account. However, a serious problem may occur due to the high computational time and space complexity of rule bases used to describe the model with multiple input variables with proper accuracy. The "exponential explosion" allows little general systems application or real time application of the classical fuzzy algorithms where the number of input variables becomes large [3]. For k input variables, and where T fuzzy linguistic terms are required in every dimension of X for all the á-covers, the number of fuzzy rules covering X at least to á is

$$|R| = O(T^k) \quad (1),$$

which could be very high, unless k is very small. This "exponential explosion" in the number of rules is a major problem hindering the application of fuzzy techniques beyond the area of fuzzy control systems. Besides, the time required in searching the large rule base in order to perform inference might become impractical for any potential applications.

There are basically three possible solutions in general to overcome this problem. First is to reduce the value of T, this can be performed by fuzzy rule interpolation techniques [4, 5, 6, 7]. Secondly, reducing the value of k; which is more effective. In 1991, Sugeno et al. [8] had proposed the use of the hierarchical structured rule bases, through little subsequent application of this method has been done as their method required significant manual crafting. The third possible solution is to decrease both T and k simultaneously has been examined in Kóczy et al. [3].

The purpose of this paper is to look at the third solution of simultaneously decreasing both T and k. We present a hierarchical fuzzy rule base algorithm that extracts its rules directly from the measured or simulated data. The hierarchical fuzzy system is used to perform classification with the purpose of reducing the complexity of computation for inference.

II. Construction of Hierarchical Fuzzy System

The first step in constructing any fuzzy rule base for performing any reasonable control and modelling in classification is to perform necessary preprocessing. The nature of the model as well as the boundary of the domains for modelling have to be determined, as this will enable a set of data to be measured or simulated in order to cover the whole universe of discourse in the fuzzy control domain. Another crucial step in preprocessing is to remove redundant inputs from the state space. Sugeno et al. [8] used a number of correlative measures to perform this. Other techniques that make use of neural networks can also be used [9, 10, 11]. In this paper, we assume all the preprocessing stages have been performed prior to the building of the hierarchical fuzzy system. The set of available numerical data can be used directly by the following proposed algorithm.

A. *Converting Numerical Values to Fuzzy Rules*

In most fuzzy control applications, the measured or simulated data is the basic knowledge available to build the fuzzy control rule base. In our hierarchical fuzzy system, the first step is to translate all the available data into linguistic fuzzy rules using linguistic labels. The following algorithm outlines the steps in extracting the fuzzy linguistic rules from the available input-output pairs.

For k inputs, the given input-output data pairs with n patterns are:

$$
\begin{gathered}
(x_1^1, x_2^1, x_3^1, \ldots, x_k^1; y^1) \\
(x_1^2, x_2^2, x_3^2, \ldots, x_k^2; y^2) \\
\vdots \\
(x_1^n, x_2^n, x_3^n, \ldots, x_k^n; y^n)
\end{gathered}
$$

The number of linguistics terms T required and the distribution of the data in each dimension of the domain are determined. For ease of interpretation and computational simplicity, the shape of the membership functions used in this algorithm is triangular. In this case, we will obtain for every $x \in X$,

$$A_t : X \rightarrow [0,1] \qquad (2)$$

After the fuzzy regions and membership functions have been distributed, the available input-output pairs will be mapped. If the value cuts on more than one membership function, the one with the maximum membership grade will be assigned to the value:

$$R_n \Rightarrow [x_1^n(A_{11}, \max), x_2^n(A_{12}, \max), x_3^n(A_{13}, \max), \ldots, x_k^n(A_{1k}, \max); y^n(B_1, \max)] \qquad (3)$$

After all the input-output values have been assigned a fuzzy linguistic label, Mamdani type fuzzy rules are then formed.

After fuzzy rules have been generated from each data point, repetitive rules are removed. Depending on the quality of the simulation or experiment set up, noise may create some conflicting fuzzy rules. In this case, the number of repetitions of the fuzzy rules and the firing strengths of the rules will be examined to resolve the conflict. Normally, with appropriate preprocessing, the total number of rules generated in this stage will follow the formulae presented in equation (1).

B. *Level Identification*

After obtaining the fuzzy rules that can represent the set of available data, the next step is to build the hierarchy levels in the hierarchical fuzzy system. Our notion is to allow one input state variable to be used in one level of the hierarchy, which structures the search space for the subsequent steps. The problem arises as in which input variable should be placed at the top and each level of the hierarchy. Our proposed algorithm is as follows.

The next step involves the identification of the contribution from each input. If we build the hierarchical fuzzy system from the most important input at the top to the least important are at the bottom, this will assist the pruning stage as will be discussed later.

In this paper, the regularity criterion (RC) [12] is used because it was at hand. However, other input contribution measures are also suitable for this purpose.

Let us assume that the C_i of all the inputs are re-arranged from lowest RC value to highest RC value as follows:

$$[C_1, C_2, C_3, \ldots, C_k] \qquad (8)$$

We will start by building from the top of the hierarchy with input, x_{C_i}, which has the lowest RC value. Let us assume that this input x_{C_i} has T linguistics terms, therefore this input can has T branches of sub rules that can be formed in the next level down the hierarchy. If we have two input variables and one output variable, and all variables have T linguistic terms, our hierarchical fuzzy rule base will have the following levels of rules:

R1:

- If x_1 is A_{11} then use R21
- If x_1 is A_{12} then use R22
- ⋮
- If x_1 is A_{1T} then use R2T

R21:

- If x_2 is A_{21} then y is B_{211}
- If x_2 is A_{22} then y is B_{212}
- ⋮
- If x_2 is A_{2T} then y is B_{21T}

⋮

R2T:

- If x_2 is A_{21} then y is B_{2T1}
- If x_2 is A_{22} then y is B_{2T2}
- ⋮
- If x_2 is A_{2T} then y is B_{2T1}

C. *Pruning Stage*

After the hierarchical fuzzy rule base has been constructed, we can prune the hierarchy to reduce the complexity of the fuzzy rule base. This will improve the inference efficiency of the fuzzy system. Without the use of the RC method, we would have to try *k*(k-1)* ways of building the hierarchical fuzzy system. Then, we would have to decide which one would give us the best hierarchical fuzzy rule base depending on how much it can be pruned.

Our proposition here is that the probability of having the input with the lowest RC value at the top of the hierarchy to the input with the highest RC value at the bottom will effectively allow more sub-rule bases to be pruned. The contribution of the inputs for predicting the classification output is a measure of how well the output classification correlates to the input space. The more correlated the output classification to the input variable, the better chance that the classification can be determined without moving down the sub-rules in the hierarchical fuzzy system.

Pruning is performed by examining the hierarchy from the bottom to the top of the hierarchical fuzzy system using two measures:

(1) if the output classification is the same for the whole sub-rule base, it can be pruned by moving the classification label up the hierarchy, or

(2) if the output classification can be interpolated by two neighbouring fuzzy rules, it is also removed from the sub-rule base.

Taking the following hierarchical fuzzy system as illustration.

R1:
If Experience is Little then use R21
If Experience is Some then use R22
If Experience is Good then use R23
R21:
If Age is Young then Salary is Basic *
If Age is Middle then Salary is Basic *
If Age is Old then Salary is Basic *
R22:
If Age is Young then Salary is Basic
If Age is Middle then Salary is Fair **
If Age is Old then Salary is High
R23:
If Age is Young then Salary is High *
If Age is Middle then Salary is High *
If Age is Old then Salary is High *

Based on the two factors proposed, the hierarchical fuzzy rule base can be reduced. Note that the sub-rule with ** in R22 can be interpolated using any fuzzy interpolation technique, therefore it can also be removed from the rule base.

R1:
If Experience is Little then Salary is Basic
If Experience is Some then use R22
If Experience is Good then Salary is High
R22:
If Age is Young then Salary is Basic
If Age is Old then Salary is High

D. *Classification Inference*

As any input value that fed into the constructed hierarchical fuzzy system used for control will normally cut more than one membership functions in the top level, in this case, the one with the maximum membership grade will be allowed to propagate to the next level for classification. As only the classification label is of interest and the fuzziness of the value is resolved in each level, no defuzzification technique is necessary at the lower end of the hierarchy.

III. CASE STUDY AND DISCUSSIONS

A case study in determining the salary category based on age, experience and contacts are used to demonstrate the use of the proposed hierarchical fuzzy system. A simulated case is used to generate a total of 200 data points, to be used as the set of input-output pairs to construct the rule base. The proposed algorithm is then used to construct the hierarchical fuzzy rule base.

After the first stage of converting all the numerical values to fuzzy rules, we have generated the following fuzzy rules using membership functions as shown in Figure 1.

Age (Age) = {Young (Y), Middle (M), Old (O)};
Experience (Exp) = {Little (L), Some (S), Good (G)};
Contacts (Con) = {Poor (P), Normal (N), Quality (Q)};
Salary (Sal) = {Basic (B), Fair (F), High (H)}

R1: If Age is Y and Exp is L and Con is P then Sal is B
R2: If Age is Y and Exp is L and Con is N then Sal is B
R3: If Age is Y and Exp is L and Con is Q then Sal is B
R4: If Age is Y and Exp is S and Con is P then Sal is B
R5: If Age is Y and Exp is S and Con is N then Sal is B
R6: If Age is Y and Exp is S and Con is Q then Sal is F
R7: If Age is Y and Exp is G and Con is P then Sal is H
R8: If Age is Y and Exp is G and Con is N then Sal is H
R9: If Age is Y and Exp is G and Con is Q then Sal is H
R10: If Age is M and Exp is L and Con is P then Sal is B
R11: If Age is M and Exp is L and Con is N then Sal is B
R12: If Age is M and Exp is L and Con is Q then Sal is F
R13: If Age is M and Exp is S and Con is P then Sal is F
R14: If Age is M and Exp is S and Con is N then Sal is F
R15: If Age is M and Exp is S and Con is Q then Sal is F
R16: If Age is M and Exp is G and Con is P then Sal is H
R17: If Age is M and Exp is G and Con is N then Sal is H
R18: If Age is M and Exp is G and Con is Q then Sal is H
R19: If Age is O and Exp is L and Con is P then Sal is B
R20: If Age is O and Exp is L and Con is N then Sal is F
R21: If Age is O and Exp is L and Con is Q then Sal is H
R22: If Age is O and Exp is S and Con is P then Sal is F
R23: If Age is O and Exp is S and Con is N then Sal is F
R24: If Age is O and Exp is S and Con is Q then Sal is H
R25: If Age is O and Exp is G and Con is P then Sal is H
R26: If Age is O and Exp is G and Con is N then Sal is H
R27: If Age is O and Exp is G and Con is Q then Sal is H

RC measure is now performed on the data set to determine the level of each input in the hierarchy. The RC values of the three input variables are Age (338.584), Experience (149.102), and Contacts (495.492). Following the values of the RC measure, the arrangement of the level of hierarchy is Experience at the top, followed by Age and then Contacts. After the hierarchy has been established, pruning is performed based on the two factors proposed earlier. The following page shows the final hierarchical fuzzy rule base after pruning. Effectively, the total number of fuzzy rules in this hierarchical fuzzy rule base is 15, which is a notable reduction from the original 27 fuzzy rules.

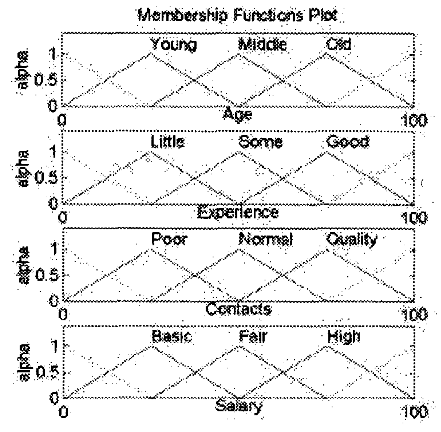

Figure 1: Membership functions for all the variables.

R1:
- If Exp is L then use R21
- If Exp is S then use R22
- If Exp is G then Sal is H

R21:
- If Age is Y then Sal is B
- If Age is M then use R32
- If Age is O then use R33

R22:
- If Age is Y then use R32
- If Age is M then Sal is F
- If Age is O then use R36

R32:
- If Con is Q then Sal is F
- If Con is NOT Q then Sal is B

R33:
- If Con is P then Sal is B
- If Con is Q then Sal is H

R36:
- If Con is Q then Sal is H
- If Con is NOT Q then Sal is F

IV. Conclusion

In this paper, we have proposed a hierarchical fuzzy system that can be used for classification in fuzzy control applications. The proposed algorithm for building the hierarchical fuzzy rule base from given input-output pairs has been shown to be effective in reducing the size of the rule base. The *k* inputs should be reduced in any preprocessing stage by removing any redundant inputs before using this hierarchical fuzzy algorithm. Further, in this proposed algorithm, the inputs are separated into different levels which also modify the effect of *k*. As fuzzy interpolation is to be used, in the pruning stage of the hierarchical fuzzy rule base the *T* is also reduced to allow better computation efficiency. In this preliminary research the hierarchical fuzzy system can only handle classification applications, further work will be needed to extend this hierarchical fuzzy system for applications that require numeric values as the output.

V. Acknowledgement

This research is supported by the Australian Research Council and Domonkos Tikk is funded by the Hungarian Scientific Research Fund Grant No. D034614

References

1. Zadeh, L.A. "Outline of a new approach to the analysis of complex systems and decision processes," IEEE Transactions on System, Man and Cybernetics (1973) 28-44.
2. Mamdani, E.H. and Assilian, S "An experiment in linguistic synthesis with a fuzzy logic contoller," International Journal of Man-Machine Studies (1975) 1-13.
3. Kóczy, L.T., Hirota, K. and Muresan, L. "Interpolation in Hierarchical Fuzzy Rule Bases," Proceedings of the 9th IEEE International Conference on Fuzzy Systems (2000) 471-477.
4. Kóczy, L.T., and Hirota, K. "Approximate reasoning by linear rule interpolation and general approximation," International Journal of Approx. Reason, vol. 9 (1993) 197-225.
5. Gedeon, T.D., and Kóczy, L.T. "Fuzzy rule interpolation using an additive conservative strategy," Proceedings of 5th European Congress on Intelligent Techniques and Soft Computing (EUFIT'97), (1997) 910-913.
6. Gedeon, T.D., and Kóczy, L.T. "Conservation of fuzziness in rule interpolation," Intelligent Technologies, vol. 1 International Symposium on New Trends in Control of Large Scale Systems (1996) 13-19.
7. Wong, K.W., Gedeon, T.D., and Tikk, D. "An improved multidimensional alpha-cut based fuzzy interpolation technique," Proceedings of Artificial Intelligence in Science and Technology (2000) 33-38.
8. Sugeno, M., Murofushi, T., Nishino, J., and Miwa, H. "Helicopter flight control based on fuzzy logic," Proceedings of Fuzzy Engineering toward Human Friendly Systems '91 (1991) 1120-1124.
9. Gedeon, T.D, "Indicators of input contributions: comparing functional measures," Proceedings of IEEE Region Ten Conference (TENCON) – Digital Signal Processing Applications (1996) 520-525.
10. Gedeon, T.D., "Data mining of inputs: analysing magnitude and functional measure," International Journal of Neural Systems, vol. 8 (2) (1997) 209-218.
11. Fung, C.C., Wong, K.W., and Crocker, H. "Determining Input Contributions for a Neural Network Based Porosity Prediction Model," Proceedings of 8th Australian Conference on Neural Network (ACNN'97) (1997) 35-39.
12. Sugeno, M., and Yasukawa, T. "A Fuzzy-Logic-Based Approach to Qualitative Modeling," IEEE Tansactions on Fuzzy Systems, vol. 1, no. 1 (1993) 7-31.