EFFECTIVE FUZZY SYSTEMS FOR COMPLEX STRUCTURED DATA

T.D. GEDEON[1], L.T. KÓCZY[1,2], K.W. WONG[1] and P. LIU[1]

[1] *School of Information Technology*
Murdoch University
Perth, Australia

[2] *Dept. of Telecom. and Telematics*
Budapest Univ. Tech. and Economics
Budapest, Hungary

Abstract

The tendency and main motivation of various Artificial and Computational Intelligence approaches has been to cope with very complex and often analytically unknown systems, in the sense of identifying approximate models, controlling or generating decision support for them. The first major step was symbolic rule based expert systems, which had the disadvantage of not utilising the internal structure of the problem space (such as ordering / partial ordering and metrics or similarity within the space). Subsymbolic approaches built in these specific mathematical features, for example by the fuzzy membership function. These approaches still have not successfully tackled problems with reasonably high numbers of input variables because of the high computational cost involved, nor have they solved the problem of dealing with problems with complex and interdependent features or where data is missing.

Keywords:

Fuzzy systems, fuzzy signatures, complex features

Introduction

Fuzzy control and decision support systems are still the most important applications of fuzzy theory [1,2]. This is a general form of expert control using fuzzy sets representing vague / linguistic predicates, modelling a system by If ... then rules. In the classical approaches of Zadeh [3] and Mamdani [4], the essential idea is that an observation will match partially with one or several rules in the model, and the conclusion is calculated by evaluation of the degree of these matches and by the use of the matched rules.

A serious problem is caused by the high computational time and space complexity of rule bases describing systems with multiple inputs with proper accuracy. The complexity allows little general systems application (or real time control application) of classical fuzzy algorithms, where the inputs exceed about 6 to 10. These traditional fuzzy systems deal with very simple structured data, where the number of inputs is well defined, and values for each input occur for most or all data items. This further reduces their general applicability.

Traditional unstructured dense fuzzy rule bases form fuzzy covers of their input universe, in the sense of $\forall x \in X : \exists i \in \{i = 1...r\} : \mu_{A_i}(x) \ge t$, where r is the number of rules of the type $R_i = \text{If } x \text{ is } A_i \text{ then } y \text{ is } B_i$ and usually $t \ge 0.5$, and μ_{A_i} is the membership function of the ith antecedent set. So it is true that for every instance of an observation that there is at least one rule in the rule base for which the antecedent is “more true than false”. Neither an even cover, nor the use of symmetrical triangular membership functions is required. In practice, smooth bell or ‘S’ shaped functions are always substituted by piecewise linear membership function shapes, in most cases by trapezoidal or even triangular membership functions.

Given k inputs, and at most T linguistic terms per dimension of X for the α-cover, the number of rules covering X at least to α is $|R| = O(T^k)$ which is very high, unless k is very small. The exponential explosion in the number of rules is a major problem hindering the application of fuzzy techniques beyond the area of fuzzy control systems. We have previously developed and adapted techniques which partially address the decreasing of T, and of k (by fuzzy rule interpolation and hierarchical rule bases respectively) [5,6,7,8]. The general problem of extracting hierarchical rule base structure is probably NP-hard, requiring approximate approaches to achieve tractability. The complexity is the main problem for the general use of such reductions, and hence hindering the general application of fuzzy rule based techniques. This is a problem we are working on elsewhere.

There are many areas where objects with complex and sometimes interdependent features are to be classified and similarities / dissimilarities evaluated. Often, human experts can and must make decisions based on comparisons of cases with different numbers of data components, with even some components missing. To create effective fuzzy systems which can handle such cases, we use the novel approach of fuzzy signatures which structures data into vectors of fuzzy values, each of which can be a further vector. This tree structure is a generalisation of fuzzy sets and vector valued fuzzy sets in a way modelling the human approach to complex problems.

Fuzzy Signatures

The original definition of fuzzy sets was $A : X \to [0,1]$, and was soon extended to *L-fuzzy sets* by Goguen [9],

$$A_S : X \to [a_i]_{i=1}^{k}, a_i = \begin{cases} [0,1] \\ [a_{ij}]_{j=1}^{k_i} \end{cases}, a_{ij} = \begin{cases} [0,1] \\ [a_{ijl}]_{l=1}^{k_{ij}} \end{cases},$$

$A_L : X \to L$, L being an arbitrary algebraic lattice. A practical special case, *Vector Valued Fuzzy Sets* was introduced by Kóczy [10], where $A_{V,k} : X \to [0,1]^k$, and the range of membership values was the lattice of *k*-dimensional vectors with components in the unit interval. A further generalisation of this concept is the introduction of fuzzy signatures and signature sets, where each vector component is possibly another nested vector (right).

Each signature corresponds to a nested vector structure or, equivalently, to a tree graph. The internal structure of the signature indicates the semantic and logical connection of state variables, corresponding to the leaves of the signature graph. For example, the signature structure (right) indicates that the nine variables ($x_1,...,x_9$) describing the given problem belong to tightly connected subgroups $\{x_7, x_8\}$; $\{x_1, x_2\}, \{x_4, x_5, x_6\}, \{x_7, x_8, x_9\}$. (The subscripts correspond to the sequential positions according to the signature.) In many applications not all data can be described by full signatures. If the above structure is the prototype with full information, there might be a datum that has less, for example only as shown (left), meaning that there is only a single aggregated value available for $\{x_1, x_2\}$ and $\{x_7, x_8\}$. It is necessary that operations on signatures, even for signatures with partly different structures can be applied. The first phase of our work is the investigation of properties and operations of fuzzy signatures and signature sets, including aggregations attached to groups of variables indicated by vectorial sub-components in the prototype signature structure. Further the investigation of specific operations on data with different sub-tree signature structures will be performed.

$$\begin{bmatrix} [a_1] \\ a_2 \\ \begin{bmatrix} a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} \\ \begin{bmatrix} a_{41} \\ a_{42} \end{bmatrix} \end{bmatrix}$$

The classic approach in this area is the one by Sugeno and Yasukawa [11], which is based on the Fuzzy C-means Algorithm (FCM) of Bezdek [12]. This algorithm is suitable to detect and approximate the shape and extent of fuzzy clusters in the output, corresponding to single or multiple rules. The critical question is here, how to separate different input clusters, which merge in their output projection (“then-part”). The method suggested in that paper is not clearly executable in the general case. Regions with too complex signatures can be eliminated if they are "surrounded" by regions of lower complexity. Proximity for signatures must be defined properly, based on partial ordering of signatures. As the starting point, (partial) ordering of fuzzy sets and vectors can be considered [13].

Applications

Let S_{S_0} denote the set of all fuzzy signatures whose structure graphs are sub-trees of the structural (“stretching”) tree of a given signature S_0 . Then the signature sets introduced on S_{S_0} are defined by

$$A_{S_0} : X \to S_{S_0} .$$

In this case the prototype structure S_0 describes the “maximal” signature type that can be assumed by any element of X in the sense that any structural graph obtained by a set of repeated omissions of leaves from the original tree of S_0 might be the tree stretching the signature of some A_{S_0} .

$$\begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} \\ a_2 \\ \begin{bmatrix} a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} \\ \begin{bmatrix} \begin{bmatrix} a_{411} \\ a_{412} \end{bmatrix} \\ a_{42} \end{bmatrix} \end{bmatrix}$$

An example for the usefulness of this definition is given below. Let us think about some patients, whose daily symptom signatures are based on doctors’ assessments according to the following scheme:

$$A_S = \begin{bmatrix} \text{fever} \begin{bmatrix} 8\text{ a.m.} \\ 12\text{ p.m.} \\ 4\text{ p.m.} \\ 8\text{ p.m.} \end{bmatrix} \\ \text{blood pressure} \begin{bmatrix} \text{systolic} \\ \text{diastolic} \end{bmatrix} \\ \text{nausea} \\ \text{abdominal pain} \end{bmatrix} .$$

Let us take a few examples with linguistic values and numerical signatures:

$$A_1 = \begin{bmatrix} \begin{bmatrix} \text{none} \\ \text{none} \\ \text{slight} \\ \text{slight} \end{bmatrix} \\ \begin{bmatrix} \text{normal} \\ \varnothing \end{bmatrix} \\ \text{slight} \\ \text{slight} \end{bmatrix} \mapsto \begin{bmatrix} \begin{bmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.2 \end{bmatrix} \\ \begin{bmatrix} 0.5 \\ \varnothing \end{bmatrix} \\ 0.25 \\ 0.25 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \begin{bmatrix} \varnothing \\ \varnothing \\ \text{moderate} \\ \text{moderate} \end{bmatrix} \\ \begin{bmatrix} \text{slightly high} \\ \text{rather high} \end{bmatrix} \\ \text{slight} \\ \text{none} \end{bmatrix} \mapsto \begin{bmatrix} \begin{bmatrix} \varnothing \\ \varnothing \\ 0.4 \\ 0.4 \end{bmatrix} \\ \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix} \\ 0.25 \\ 0.0 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} \begin{bmatrix} \text{rather high} \\ \text{high} \\ \text{rather high} \\ \text{rather high} \end{bmatrix} \\ \begin{bmatrix} \text{rather high} \\ \text{very high} \end{bmatrix} \\ \text{none} \\ \varnothing \end{bmatrix} \mapsto \begin{bmatrix} \begin{bmatrix} 0.8 \\ 0.6 \\ 0.8 \\ 0.8 \end{bmatrix} \\ \begin{bmatrix} 0.8 \\ 1.0 \end{bmatrix} \\ 0.0 \\ \varnothing \end{bmatrix}$$

Of course, normally the blood pressure values would initially rather be expressed by the physician as e.g. 75 / 120, which could then be converted to the linguistic values as appropriate for the patient, taking into account contextual information such as the higher normal blood pressure of infants and children and so on. As for most techniques, there is a significant role for the use of background knowledge of domain experts in data preprocessing.

Note that the structures above are different, which is the point as much real world data is like this. For patient 2, we have only 2 measurements for fever. The structure of the fuzzy signature contains some information by the association of vector components. The use of aggregation operators allows us to compare components irregardless of the different numbers of sub-components. Such aggregation operators would in general be designed for each vectorial component with the assistance of a domain expert. In this case, let us assume that the time of day of fever is less significant, and that the maximum value is most important. (In this assumption, the spacing of measurements must therefore be to ensure a reasonable coverage of the data source.) The three will be reduced to the following forms:

$$A_{1f} = \begin{bmatrix} 0.2 \\ \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \\ 0.25 \\ 0.25 \end{bmatrix}, A_{2f} = \begin{bmatrix} 0.4 \\ \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix} \\ 0.25 \\ 0.0 \end{bmatrix}, A_{3f} = \begin{bmatrix} 0.8 \\ \begin{bmatrix} 0.8 \\ 1.0 \end{bmatrix} \\ 0.0 \\ \varnothing \end{bmatrix}$$

The “fever component” can be verbally rewritten as “slight”, “moderate” and “rather high”, respectively. The signatures above still contain sufficient information about the “worst case fever” of each patient, while the detailed knowledge on the daily tendency of the fever is lost. This hierarchically structured access to the information is a key benefit of the fuzzy signatures.

We could continue this process further completely, and determine an overall “abnormal condition” measure $A_{1o} = [0.25], A_{2o} = [0.4], A_{3o} = [1.0]$.

Conclusion

We have described a technique for dealing with problems with complex and interdependent features or where data is missing. This was by the notion of fuzzy signatures, which extends the concept of vectorial fuzzy sets to components with varying numbers of components. We have demonstrated with an example the benefit of this hierarchical structuring of data. This hierarchical structuring allows the further use of domain experts as the information can be abstracted to higher levels analogous to patterns of human expert decision making.

References

[1] Chong, A., Gedeon, T.D., Wong, K.W., "Extending the Decision Accuracy of A Bioinformatics System," in Reznik, L., Kreinovich, V. (Eds.) *Soft Computing in Measurement and Information Acquisition*, (Physica-Verlag, Springer-Verlag, in press).

[2] Gedeon, T.D. “Clustering Significant Words using their Co-occurrence in Document Sub-Collections,” *Proceedings 7th European Congress on Intelligent Techniques and Soft Computing (EUFIT’99)*, Aachen, 1999, pp. 302-306.

[3] Zadeh, LA “Outline of a new approach to the analysis of complex systems and decision processes,” *IEEE Transactions on Systems, Man and Cybernetics*, 1973, pp. 28-44.

[4] Mamdani, EH and Assilian, S “An experiment in linguistic synthesis with a fuzzy logic controller,” *Int. Journal of Man-Machine Studies*, 1975, pp. 1-13.

[5] Gedeon, T.D. and Kóczy, L.T., “Hierarchical co-occurrence relations,” *IEEE International Conference on System Man and Cybernetics (SMC’98)*, San Diego, 1998, pp. 2750-2755.

[6] Kóczy, L.T. and Hirota, K “Interpolation in structured fuzzy rule bases,” *Proc. 2nd IEEE Int. Conf. on Fuzzy Systems*, San Francisco, 1993, pp. 402-405.

[7] Kóczy, L.T. and Hirota, K “Approximate reasoning by linear rule interpolation and general approximation,” *Int. Journal of Approximate Reasoning*, 1993, no. 9, pp. 197-225.

[8] Gedeon, T.D., Kóczy, L.T., “Fuzzy rule interpolation using an additive conservative strategy,” *Proc. 5th European Conf. Intell. Tech. & Soft Comput.*, Aachen, 1997, pp. 910-913.

[9] Goguen, J. A. “L-fuzzy Sets,” *J. Math. Anal. Appl.* 18, 1967, 145-174.

[10] Kóczy, L. T. “Vectorial I-fuzzy Sets,” in: *M. M. Gupta & E. Sanchez (eds.): Approximate Reasoning in Decision Analysis*, North Holland, Amsterdam, 1982, 151-156.

[11] Sugeno, M and Yasukawa, T “A fuzzy logic based approach to qualitative modelling,” *IEEE Trans. Fuzzy Systems*, vol. 1, no. 1, Feb 1993, pp. 7-31.

[12] Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*, NY: Plenum Press, 1981.

[13] Kóczy, L.T., and Hirota, K. “Ordering, distance and closeness of fuzzy sets,” *Fuzzy Sets and Systems* 59, 1993, pp. 197-225.