Clustering Methods in Fuzzy Control

F. Klawonn, R. Kruse

Department of Computer Science, University of Braunschweig
D-38106 Braunschweig, Germany

Summary: Fuzzy controllers can be interpreted as an interpolation technique on the basis of fuzzy clusters of input/output pairs. It is therefore obvious that fuzzy clustering algorithms are a promising tool for supporting the design of a fuzzy controller when data of the process to be controlled are available.

This paper discusses the possibilities and limitations of fuzzy clustering for fuzzy control.

1. Introduction

Most of the classical control techniques are based on a mathematical or physical model of the process to be controlled, usually described by a set of difference or differential equations. These techniques are applicable when a suitable, not too complex mathematical description of the process can be provided. But there is a great number of processes with parameters for which no appropriate mathematical model is known. Nevertheless, many of these processes can be controlled by a human operator.

The idea behind fuzzy control is to model the behaviour of a human operator. Thus fuzzy controllers aim at determining a static control function on the basis of linguistic if–then–rules. The control function assigns to each tuple of measured input values (for example error and and change of error) a suitable control or output value that forces the process in the direction of the desired state. The control rules are of the form

$$\textit{If } \textit{input}_1 \textit{ is } A_1 \textit{ and } \ldots \textit{ and } \textit{input}_n \textit{ is } A_n \textit{ then output is } B \qquad (1)$$

where $A_1, \ldots, A_n, B$ are linguistic expressions of the form *approximately zero, (negative) small, (positive) big*, etc. that specify a vague description of the value of the corresponding variable. These linguistic expressions are represented by fuzzy sets.

In section 2 we will see that each fuzzy set can be interpreted as representing a crisp value in a vague environment admitting some small error. In this sense, the fuzzy sets are in some way fuzzy clusters. Section 3 gives a short review of basic fuzzy control techniques in the light of the above mentioned interpretation of fuzzy sets. Section 4 is devoted to the topic of generating fuzzy sets automatically by fuzzy clustering from a data set. It will be shown that fuzzy clustering can be useful for this task, but in many cases fuzzy clustering will not meet the requirements needed for fuzzy control.

2. Interpretation of the Fuzzy Sets

The idea behind fuzzy sets is to extend the concept of membership to a set to graded membership, i.e. not to restrict to the two membership degrees 1 ('belongs to the set') and 0 ('does not belong to the set'). Therefore, from a formal point of view, a fuzzy set μ on the domain X is a generalized characteristic function $\mu : X \to [0,1]$, assigning to each $x \in X$ its membership degree $\mu(x)$ to the fuzzy set μ, or, if μ is intended to represent a linguistic expression like *approximately zero*, the degree to which x satisfies the linguistic expression associated with μ.

Although it is very appealing to interpret the value $\mu(x)$ as the degree to which x belongs to the fuzzy set μ, there is a need for a concrete interpretation of membership degrees. Otherwise it is impossible to assign suitable membership degrees to elements or to compare or combine fuzzy sets specified by different people.

The approaches to the interpretation of membership degrees can be categorized into three types: uncertainty (for instance in the case when fuzzy sets are seen as possibility distributions like in Dubois and Prade (1988)), preference (as in Dubois and Prade (1993)), and similarity (see for example Trillas and Valverde (1984) or Kruse et al. (1993)).

In this paper where we only consider fuzzy control we restrict ourselves to the latter interpretation. In order to explain this interpretation of membership degrees in terms of similarity or indistinguishability, let us consider the following simple example.

Let δ be a (pseudo–)metric on X. Then

$$E_\delta : X \times X \to [0,1], \qquad (x,y) \mapsto 1 - \min\{\delta(x,y), 1\}$$

is called an equality or similarity relation. The value $E_\delta(x,y)$ reflects the similarity of x and y. Note that it might be reasonable to apply a scaling to the metric δ. For example, in the case of a discrete metric δ which maps X into the natural numbers, one could use $c \cdot \delta$ instead of δ where $0 < c < 1$ is a suitable scaling factor. Otherwise the corresponding similarity relation would simply be the crisp equality.

Given a crisp element $x_0 \in X$, one should take the similarity relation into account and consider the 'equivalence class' of x_0 with respect to E_δ, i.e. the crisp set $\{x_0\}$ is extended to the fuzzy set

$$\mu_{x_0} : X \to [0,1], \qquad x \mapsto E_\delta(x_0, x)$$

containing all elements that are similar to x_0. In this way, under consideration of the similarity relation E_δ or its dual concept, the metric δ, each crisp value x_0 induces a fuzzy set or fuzzy cluster μ_{x_0}. It is very important to note that in the case $X = \mathbb{R}$ and $\delta(x,y) = |x-y|$ the fuzzy set μ_{x_0} has a triangular shape like those fuzzy sets that are commonly used in fuzzy control.

Of course, the fuzzy sets appearing in fuzzy control are in general not so simple that they can be considered to be of the form μ_{x_0} with the standard metric on the real numbers as the underlying metric. Usually a transformation of the real line is assumed. The transformation is induced by a scaling function $c : \mathbb{R} \to [0, \infty[$ that assigns to each $x \in \mathbb{R}$ a scaling factor $c(x) \geq 0$. The greater this scaling factor is, the stronger is the distinguishability of values in the neighbourhood of x. Therefore, the underlying metric is given by the following formula.

$$\delta_c(x, y) \quad = \quad \left| \int_x^y c(s)ds \right| .$$

Note that generally in fuzzy control neither the scaling function c nor the the metric δ_c or its corresponding similarity relation E_{δ_c} is explictly considered. But it was shown by Klawonn and Kruse (1993) that in most cases an appropriate metric or even a scaling function can be found so that the fuzzy sets can be interpreted in the above mentioned way. The problem of finding such a metric or scaling function for a given fuzzy partition, i.e. a family of fuzzy sets is discussed by Höhle and Klawonn (1992) and Klawonn (1994).

Before we can relate these ideas to fuzzy control, we have to mention the problem of aggregating similarity relations on different domains. Let δ_i $(i = 1, 2)$ be a metric on X_i and let E_{δ_i} denote its corresponding similarity relation. When we consider the product space $X_1 \times X_2$ we have various possibilities to derive a similarity relation on $X_1 \times X_2$ from E_{δ_1} and E_{δ_2}. For reasons of simplicity, we restrict ourselves to

$$E : (X_1 \times X_2) \times (X_1 \times X_2) \to [0, 1],$$

$$((x_1, x_2), (y_1, y_2)) \quad \mapsto \quad \min\{E_{\delta_1}(x_1, y_1), E_{\delta_2}(x_2, y_2)\}.$$

Note that E is induced by the metric

$$\delta((x_1, x_2), (y_1, y_2)) \quad = \quad \max\{\delta_1(x_1, y_1), \delta_2(x_2, y_2)\}$$

and that δ is an ultrametric if δ_1 and δ_2 are ultrametrics.

3. Fuzzy Control

Fuzzy controllers are used to describe static control functions that assign to each tuple of measured input values of a process a suitable output value for the control variable that forces the system in the direction of the desired state. In order to achieve this, for each input variable and for the output variable a fuzzy partition of the corresponding domain is specified, i.e. a family of fuzzy sets for each domain has to be determined. A typical fuzzy partition is illustrated in figure 1.

Each of the fuzzy sets is associated with a linguistic expression like *approximately zero, (negative) small, (positive) big,* etc. These linguistic expressions are incorporated in the control rules of the form (1) mentioned in the

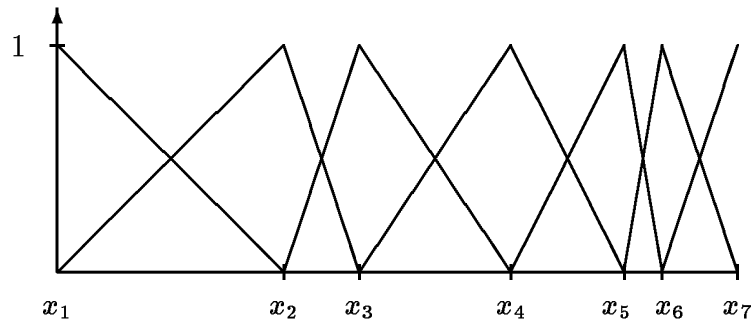

Figure 1: A typical fuzzy partition.

introduction.

For a given tuple $(\xi_1, \ldots, \xi_n)$ of measured inputs for each rule the 'matching degree' is computed, i.e. if the fuzzy set μ_i is associated with the linguistic expression A_i, then the matching degree of rule (1) for the input $(\xi_1, \ldots, \xi_n)$ is

$$\alpha_1 = \min\{\mu_1(\xi_1), \ldots, \mu_n(\xi_n)\}. \tag{2}$$

The matching degree determines how strong the rule is applicable and effects the output of the rule. Typically, the output of one rule is defined as the fuzzy set $\min\{\mu, \alpha_1\}$ where μ is the fuzzy set associated with the linguistic expression B and α_1 is the value defined in equation (2). The output fuzzy set of the system of rules is obtained by aggregating the output fuzzy sets of the single rules by the maximum. In order to get a crisp output value this fuzzy set has to be 'defuzzified' which is often done by the center–of gravity–method that takes the value under the center of gravity of the fuzzy set as output.

It would lead us to far to discuss here the details of the computations carried out in a fuzzy controller and we refer for an overview to Kruse et al. (1993). What is interesting from the viewpoint of cluster analysis is the following.

The fuzzy sets can be interpreted as fuzzy clusters, i.e. as representing crisp values or prototypes with respect to an underlying similarity relation. If the fuzzy set μ_i represents the value ξ_i, then the rule (1) specifies the output for the input tuple $(\xi_1, \ldots, \xi_n)$. Therefore, the rule base determines a partial function that assigns an output value to some input tuples. In this sense, fuzzy control can be interpreted as an interpolation method in the presence of indistinguishability characterized by similarity relations. For details see Klawonn and Kruse (1993).

Note that for each input domain and for the output domain we have a specific similarity relation. Taking equation (2) into account, that determines how the matching degree of an input tuple is calculated for one rule, one can show that the similarity relation on the different domains are aggregated to a similarity relation on their product space as proposed at the end of the previous section.

4. Fuzzy Clustering and Fuzzy Control

As we have explained in the previous section, the main ingredients for a fuzzy controller are fuzzy partitions of the output and input domains and a rule base. The fuzzy sets of the fuzzy partitions can be interpreted as crisp values in the presence of indistinguishability or as fuzzy clusters. The rule base corresponds to a partial function.

Although it is intuitively appealing to formulate the knowledge about the operator's behaviour in the form of if–then–rules involving linguistic expressions like *approximately zero* and to model these expressions with fuzzy sets, it is often very difficult to specify a suitable rule base and to determine appropriate fuzzy sets. Even if the principal shape of the fuzzy sets is known, it is not clear how to choose the exact values for the membership degrees. Thus an automatic generation of the fuzzy partitions and the rule base from data gained from observing the process and the operator's behaviour is desirable.

A fuzzy controller constructed in this way can be used to simulate the operator's behaviour for automatic control. But it can also be the basis of understanding the operator's control strategy in order to implement an improved strategy in a fuzzy controller with modified fuzzy partitions and a modified rule base.

It is near at hand to think of applying standard fuzzy clustering techniques like the fuzzy c–means algorithm. For a description of the fuzzy c–means algorithm see for example Bezdek (1973), Bezdek and Pal (1992), or Dunn (1974).

The problem is that in most cases the data are better suited for regression than for clustering as illustrated in figure 2, since the operator's actions are often more or less continuous distributed over the input domains.

There are, of course, exceptions, especially for chemical processes where the operator's action consists in adding half a ton of some liquid to the process. Such drastic actions are usually carried out, when the operator observes that a certain value exceeds or falls below some threshold. In this case, one may obtain data that have the characteristics of those in the left side of figure 2.

One might ask the question, why we insist on a fuzzy controller, when a regression technique might be applied easily. The reason is that although it might be very simple to obtain a suitable control function by such a technique, this function is difficult to interpret. This means, tuning and adjustments to changing parameters of the process are more or less impossible using the regression function. Since the behaviour of a fuzzy controller becomes clear from its fuzzy sets and its rule base, such changes can be carried out, when the control function is determined by a fuzzy controller.

Figure 2: Data that are better suited for regression (left side) and for clustering (right side).

Therefore, usually some heuristic filtering has to be applied to the data that extracts data in the form of those on the right side of figure 2 from data in the form of those on the left side.

Another problem is caused by the fact that for each input and output domain a separate fuzzy partition is needed in order to formulate the control rules in the form of (1). However, the data are in general observed in the product space of all input spaces and the output space. Projecting the data before applying fuzzy clustering will lead to severe problems, since (fuzzy) clusters that do not interfere in the product space might melt together when projected to some one–dimensional space so that the fuzzy clustering of the projected data comes up with clusters that cannot be associated with (fuzzy) clusters in the product space. On the other hand, applying a fuzzy clustering algorithm in the product space and then projecting the fuzzy clusters may yield projected clusters that strongly overlap. Nevertheless, although this overlapping (i.e. fuzzy partitions with strongly overlapping fuzzy sets) is not desirable, it is still possible to construct a fuzzy controller on the basis of these fuzzy partitions, whereas the other method – first projecting and then clustering – may lead to incoherent fuzzy partitions.

Again, some suitable preprocessing of the data that forces the (fuzzy) clusters to be near the grid points in the product space, can help avoiding this problem.

Another question is, whether the fuzzy c–means algorithm is really well–suited for generating fuzzy partitions for a fuzzy controller. One of the assumption of this algorithm is that for each element, the sum of its membership degrees to all (fuzzy) clusters is equal to one. This condition is very appealing, since in fuzzy control it is very common to assume that for any element, the sum of its membership degrees to all fuzzy sets of a fuzzy partition is also one. However, in fuzzy control there is in general the stronger assumption that usually the intersection of the supports of more than two fuzzy sets of a fuzzy partition should be empty – a condition which is normally not guaranteed by the fuzzy c–means algorithm.

An alternative to the fuzzy c–means algorithm is the possibilistic approach to fuzzy clustering proposed by Krishnapuram and Keller (1993). In opposition to the fuzzy c–means algorithm the requirement that the membership degrees to all fuzzy clusters sum up to one for each element is given up in possibilistic clustering. But possibilistic clustering can tend to non–overlapping fuzzy clusters. As a consequence for the corresponding fuzzy partitions, there can be areas which are covered by no fuzzy set at all. This means that for input values falling into such areas, no control action is specified which may lead to difficulties in controlling the process.

5. Conclusions

The discussion of fuzzy clustering and fuzzy control has shown that there are strong relations between these two fields. However, a lot of problems arise, when fuzzy clustering has to be applied to fuzzy control. Some of these problems can be overcome by some suitable filtering or preprocessing of the data, but there are still unsolved problems. Nevertheless, there are promising perspectives, even if one should not expect to generate an optimal fuzzy controller from data by some fuzzy clustering technique. In any case, it is reasonable to construct a rough design of a fuzzy controller by fuzzy clustering. Improvements and tuning of such a fuzzy controller can be carried out by hand or by other learning and optimization techniques in fuzzy control (for an overview see Nauck et al. (1994)), especially since most of these techniques do not learn from scratch but rely on some more or less well working first model.

References:

BEZDEK, J.C. (1973): Fuzzy Mathematics in Pattern Classification. Ph.D. Dissertation, Appl. Math., Cornell Univ., Ithaca, NY.

BEZDEK, J.C., and PAL, S.K. (1992): Fuzzy Models for Pattern Recognition. IEEE Press, New York.

DUBOIS, D., and PRADE, H. (1988): Possibility Theory. Plenum Press, New York.

DUBOIS, D., and PRADE, H. (1993): Possibility Theory, Belief Revision and Non–Monotonic Logic. Proc. EUFIT'93, Aachen, 714–719.

DUNN, J.C. (1974): A Fuzzy Relative of the ISODATA Process and its Use in Detecting Compact Well–Separated Clusters. *Journal of Cybernetics, 3, 32–57.*

HÖHLE, U., and KLAWONN, F. (1992): Fuzzy Control und Ununterscheidbarkeit. Proc. VDE–Fachtagung Technische Anwendungen von Fuzzy–Systemen, Dortmund, 3–9.

KLAWONN, F. (1994): Fuzzy Sets and Vague Environments. *Fuzzy Sets and Systems* (to appear).

KLAWONN, F., and KRUSE, R. (1993): Equality Relations as a Basis for Fuzzy Control. *Fuzzy Sets and Systems, 54, 147–156.*

KLAWONN, F., and KRUSE, R. (1993): Fuzzy Control as Interpolation on the Basis of Equality Relations. Proc. 2nd IEEE International Conference on Fuzzy Systems 1993, IEEE, San Francisco, 1125–1130.

KRISHNAPURAM, R., and KELLER, J.M. (1993): A Possibilistic Approach to Clustering. *IEEE Transactions on Fuzzy Systems, 1, 98–110.*

KRUSE, R., GEBHARDT, J., and KLAWONN, F. (1993): Fuzzy–Systeme. Teubner Stuttgart. (English translation: Foundations of Fuzzy Systems. Wiley, Chichester, 1994).

NAUCK, D., KLAWONN, F., and KRUSE, R. (1994): Neuronale Netze und Fuzzy–Systeme: Grundlagen des Konnektionismus, Neuronaler Netze und der Kopplung mit wissensbasierten Methoden. Vieweg, Braunschweig.

TRILLAS, E., and VALVERDE, L. (1984): An Inquiry into Indistinguishability Operators. In: H.J. Skala, S. Termini, E. Trillas (eds.): *Aspects of Vagueness.* Reidel, Dordrecht, 231–256.