Categorizing classes of signals by means of fuzzy gradual rules

Sylvie Galichet, Didier Dubois, Henri Prade
Institut de Recherche en Informatique de Toulouse (IRIT)
118 route de Narbonne, 31062 Toulouse Cedex 4, France
{galichet, dubois, prade }@irit.fr

Abstract

This paper presents an approach to the approximate description of univariate real-valued functions in terms of precise or imprecise reference points and interpolation between these points. It is achieved by means of gradual rules which express that the closer the variable to the abscissa of a reference point, the closer the value of the function to the ordinate of this reference point. Gradual rules enable us to specify sophisticated gauges, under the form of connected areas, inside of which the function belonging to the class under consideration should remain. This provides a simple and efficient tool for categorizing signals. This tool can be further improved by making the gauge flexible by means of fuzzy gradual rules. This is illustrated on a benchmark example.

1 Introduction

Signal functions, such as time series, medical ECG's, arc usually viewed as analytical mappings. Then, a precise representation is often used. Even when uncertainty is dealt with, it is supposed to pervade parameters of the analytical models, leading to probabilistic or interval-based processing. The main objective of this paper is to propose an alternative to this classical type of approaches by investigating the interest of a special kind of fuzzy « if., then » rules, named gradual rules [Dubois and Prade, 1992; 1994], for developing imprecise representations. Actually, the proposed strategy relies on interpolative reasoning. By specifying the interpolate representation with gradual rules, there is no need to choose an analytic form for the interpolator and an imprecise model is directly obtained from the constraints expressed by the rules. The proposed rule-based approach is an alternative to works based on fuzzy polynomials [Lowen, 1990] or fuzzy spline interpolation, e.g. [Kawaguchi and Miyakoshi, 1998], which rely on fuzzy-valued functions and still depend on the analytical form of the interpolant.

What is supposed to be known, in a precise or in an imprecise way, is the behaviour of the function at some reference points, the problem being to interpolate between these points. Figure 1 illustrates our purpose of building an imprecise model in a case where the points on which interpolation is based are imprecise and modelled by rectangular areas.

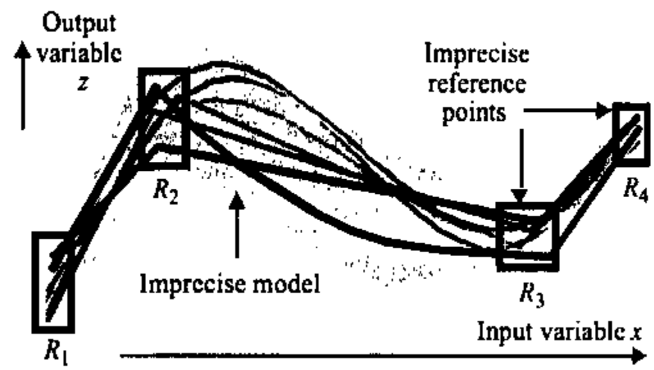


figure 1: Imprecise interpolation

So, we are no longer looking for a function, but for a relation linking the variable to the possible value of the function. Then, this relation is represented by its graphT defined on the Cartesian product $X \times Z$ (where X is the input domain, and Z the function range). A similar approach used in automated control, recently proposed in [Sala and Albertos, 2001], also considers the building of uncertain fuzzy models in the setting of the approximation of multi-valued functions called "ambiguous functions".

The paper, after some brief background on gradual rules, discusses the design of the imprecise interpolate representation in terms of gradual rules, constrained by precise or imprecise reference points. The proposed approach is then applied to the classification of time series. Moreover, a more powerful representation framework, based on fuzzy sets of gradual rules (called fuzzy gradual rules for short) is proposed. It enables us to describe areas through which it is not completely possible that the function may go.

2 Interpolation and gradual rules

The idea of imprecise interpolation suggested above is based on constraints to be satisfied. Namely the results of the interpolation should agree with the reference points. These constraints define the graph Γ of a relation on $X \times Z$.

We first consider the case of precise interpolation points P_i with coordinates (x_i, z_i), $i = 1, \ldots, n$. Then the relationT should satisfy:

$$\Gamma(x_i, z_i) = 1,$$
$$\forall z \neq z_i \in Z, \quad \Gamma(x_i, z) = 0,$$

for $i = 1, \ldots, n$. Without any further constraint on the nature of the interpolation, we only have:

$$\forall x \neq x_i \in X, \forall z \in Z, \ \Gamma(x, z) = 1.$$

Thus each interpolation point induces the constraint "If $x = x_i$ then $z = z_i$", represented by the r $(x = x_i) \to (z = z_i)$ r e $\to$ is material implication. The relation is thus obtained as the conjunction:

$$\Gamma(x, z) = \wedge_{i = 1, \ldots, n} \ (x = x_i) \to (z = z_i). \quad (1)$$

This relation is extremely imprecise since there is no constraint at all outside the reference points. Instead of using a precise type of interpolation function, one may use fuzzy rules in order to express constraints in the vicinity of the interpolation points. The idea is to use rules of the form "the closer* is to x_i, the closer:: is to z_i". The extension of equation (1) to gradual rules provides the following expression for the graph Γ:

$$\Gamma(x, z) = \min_{i = 1, \ldots, n} \ \mu_{\text{close to } x_i}(x) \to \mu_{\text{close to } z_i}(z) \quad (2)$$

where $\to$ represents Reschcr-Gaincs implication ($a \to b = 1$ if $a \leq b$ and $a \to b = 0$ if $a > b$), and $\mu_{\text{close to } x_i}(x)$ is the degree of truth of the proposition "x is close to x_i"\

We have just to define what is meant by "close to". Let A_i denote the fuzzy set of values close to x_i. It is natural to set $\mu_{A_i}(x) = 1$ if $x = x_i$ and to assume that the membership degree to A_i decreases on each side of x_i with the distance to x_i. However A_i will not be necessarily symmetrical. The simplest solution consists in choosing triangular fuzzy sets with a support denoted by $[x_i^-, x_i^+]$. In a similar way, the closeness to z_i will be modelled by a triangular fuzzy set *B,* with modal value *Zi* and support $[z_i^-, z_i^+]$. Then the interpolation relation only depends on 4n parameters $x_i^-, x_i^+ z_i^-, z_i^+$ for n interpolation points. In order to simplify their tuning, we further assume that at most two rules are simultaneously fired at each point of the input domain, i.e. $x_i \leq x_{i+1}^-$ and $x_i^+ \leq x_{i+1}$, $i = 1, \ldots, n-1$. For increasing reference points (i.e. $x_i < x_{i+1}$ and $z_i < z_{i+1}$, $i = 1, \ldots, n-1$), it has been established in [Galichet et al., 2002] that the interpolation graph is made of connected 4-sided areas (as pictured in figure 2) when the following constraints hold between the parameters:

$$(z_i^+ - z_{i+1})(x_{i+1}^- - x_{i+1}) = (z_{i+1}^+ - z_{i+1})(x_i^+ - x_{i+1}) \quad (3)$$

$$(z_{i+1}^- - z_i)(x_i^+ - x_i) = (z_i^- - z_i)(x_{i+1}^- - x_i). \quad (4)$$

Similar relationships can be obtained when considering decreasing reference points. These constraints can be related to consistency requirements between gradual rules [Dubois ct al., 1997], ensuring the non emptiness of the image of each input point via the relation Γ.

Figure 2 pictures the interpolation graph which is obtained with three interpolation points, and thus three gradual rules whose Afs and Bi's also appear in figure 2. The partitioning of *X* is obtained by cutting the intervals $[x_i, x_{i+1}]$ into three equal parts, i.e. $x_{i+1}^- - x_i = x_i^+ - x_{i+1}^- = x_{i+1} - x_i^+$, as ed by the parameters which define the fuzzy sets A_i. The extreme values z_1^- and z_3^+ are also predefined. Lastly, the other parameters can be obtained by solving the system of equations derived from (3) and (4).

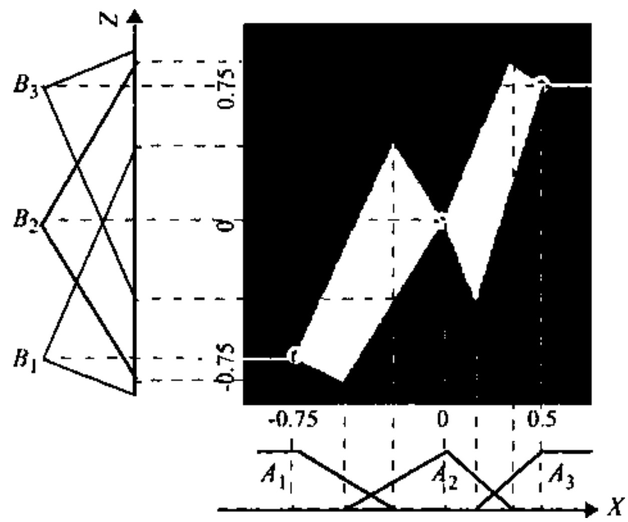


figure 2: Piecewise quadrangle-shaped graph

This approach straightforwardly extends to imprecise reference points provided that trapezoidal membership functions are used in place of triangular ones (see figure 3).

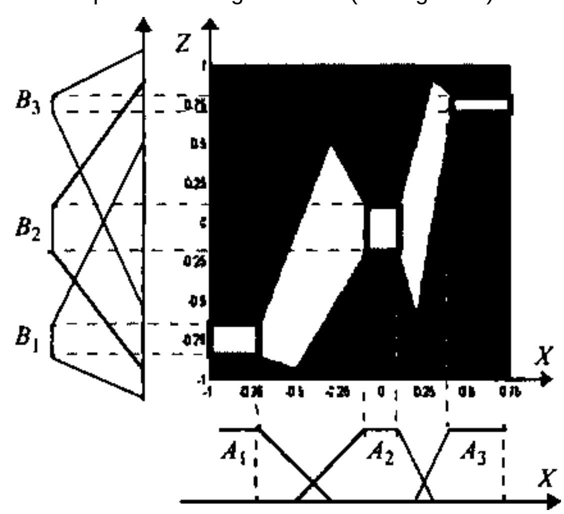


figure 3: Imprecise reference points

3 Classification of time series

Our purpose is now to illustrate how the imprecise intcrpolative representation framework can be used to classify time series, see, e.g. [Kadous, 2002]. Supervised classification is assumed, contrary to clustering techniques whose recent developments are often based on hidden Markov models [Biswas et al., 1998; Li, 2000]. The proposed experiment deals with the **«Control-Chart»** database, freely available from the UCI Data Archive [UCI KDD Archive]. It is a 6-class problem, with 100 examples of each class, a prototype of each class being presented in figure 4. A categorization of process trends, based on types of variation, is also adopted in qualitative reasoning and model-based diagnosis [Colomcr *et al.,* 2002].

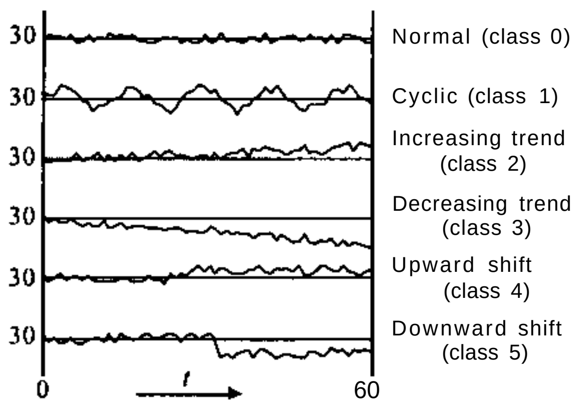


figure 4: One example of each class

Given an unlabeled time series, the aim of the classification is to decide to which class it belongs. The idea behind the proposed methodology consists in developing an imprecise model of each class. Then, the time scries to classify will be assigned to the class whose model presents the maximal adequacy with the temporal data under consideration. The imprecise models are specified using gradual rules as advocated in the previous section.

In figure 5, ten examples of class 5 are plotted simultaneously. It clearly shows that the reference points arc no longer precise. In this context, triangular membership functions are replaced by trapezoidal membership functions whose cores delimit the rectangular areas associated with the imprecise reference points. According to this slight modification, the graph plotted in figure 5 is obtained from two gradual rules, i.e. $A_1 \rightarrow B_1$ and $A_2 \rightarrow B_2$. It can be shown that the constraints on the graph shape expressed by equations (3) and (4) are still valid.

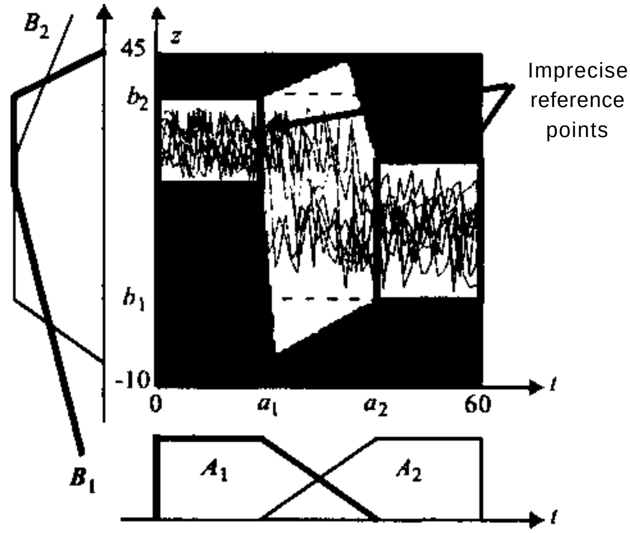


figure 5: Imprecise model of class 5

The model so built can be further improved by truncating the upper and lower parts of the quadrangle-shaped graph. An easy strategy to implement the graph cutting consists in adding a new rule that directly translates the interval-based constraint "If $t \in [a_1, a_2]$ then $z \in [b_1, b_2]$" where a_1, a_2, b_1 and b_2 arc defined in figure 5. Such an approach results in the final graph given in figure 6.

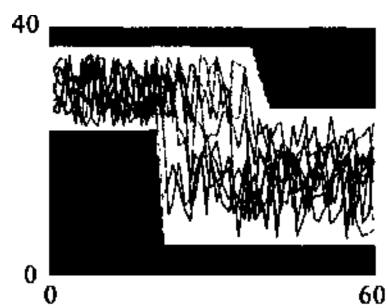


figure 6: Model of class 5 (truncated graph)

Figure 7 and 8 present the implemented models for two others classes. The first one associated with class 3 is based on reference points whose imprecision is only relative to the output. Using strong partitions with triangular input membership functions and trapezoidal output ones, imprecise linear interpolation is obtained by means of the two rules $A_1 \rightarrow B_1$ and $A_2 \rightarrow B_2$.

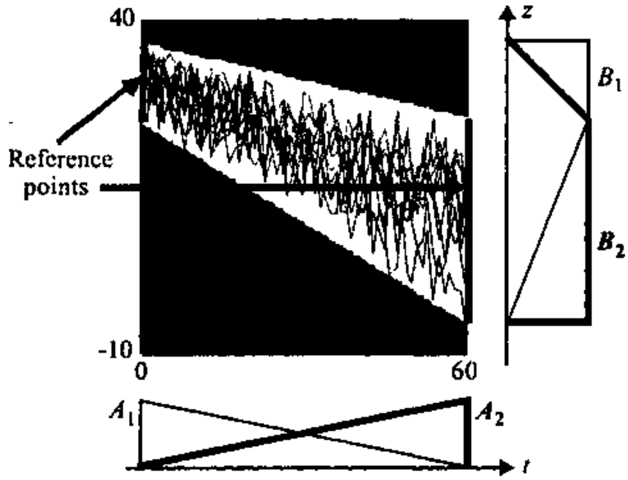
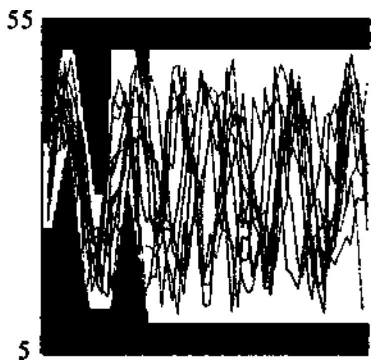


figure 7: Model for class 3 (decreasing trend)

Concerning cyclic time scries (figure 8), the non-monotonic underlying behavior induces some difficulties in the modelling process. Actually, closeness on the right and on the left of the reference points must be handled in different ways. It means that two distinct fuzzy subsets are required for correctly dealing with each reference point. In this framework, the imprecise models of figure 8 is composed of 9 gradual rules.



(c) Cyclic

figure 8: Model for class l (cyclic)

Imprecise models are built for the six classes so that the graphs include all the points that compose the training time series (10 for each class). The classification of a time-series, given as a collection of points (t_i, z_i), $t_i = i = 1, ..., 60,$ is then carried out according to its adequacy with the class models. The latter is determined from the number of points of the time series under consideration that belong to the graph of each model, that is:

$$N_i = \Sigma_{i=1,...,60} \Gamma_j(t_i, z_i), j = 1, ..., 6$$

where Γ_j denotes the model graph of the j^{th} class. The final decision is then to assign the time series, supposed unlabeled, to the class that maximizes $N_j, j = 1, ..., 6$. Applying this strategy to classify the 600 available examples, perfect classification is obtained, i.e. the error rate is null for the training examples but also for the test time scries. This result is better than the one obtained with other approaches of the same problem [Kcogh and Kasetty, 2002; Nanopoulos et al., 2001]. It is however important to be cautious about this good performance. Indeed, the discrimation between some classes is not robust. This point is illustrated by figure 9 in the case of classes 3 and 5 which are difficult to differentiate. The adequacy between the 100 time series of class 3 and models of classes 3 and 5 is plotted. It can be stated that, for many time series, the difference between both obtained scores is small, which means that a slight modification of the models would probably result in different final decisions. Actually, an important intersection between both models induces a loss of the discrimination power of the adequacy index. In this framework, one may think of improving the robustness of the classification by refining the imprecise models. One possible strategy is then to introduce some membership degrees in the four sided areas while keeping their support unchanged. This can be made by using fuzzy sets of gradual rules as shown now.

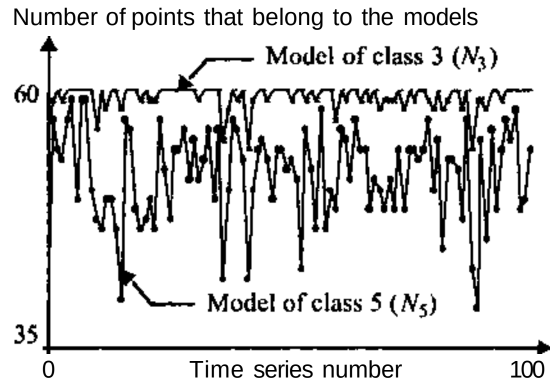


figure 9: Adequacy of class 3 examples with models of classes 3 and 5

4 Interpolate fuzzy graph

According to the previous sections, it is clear that given a set of rules, i.e. a set of reference points, a collection of crisp graphs can be obtained by varying the support parameters of the A_i's and/or th B_i's. reover, inclusion properties between the built graphs can be exhibited for controlled variations of the supports as expressed by the following statements.

P1: If $A_i \subseteq A_i^*$, $i=1, ..., n$, then $\Gamma^* \subseteq \Gamma$, where Γ and Γ^* are the graphs associated respectively with rules $A_i \rightarrow B_i$ and $A_i^* \rightarrow B_i$. Indeed, $(x, z) \in \Gamma^*$ means that $\forall i$, $A_i^*(x) \leq B_i(z)$. According to the assumption that $\forall i$, $A_i \subseteq A_i^*$, it follows that $\forall i$, $A_i(x) \leq B_i(z)$ which results in $(x, z) \in \Gamma$.

P2: If $B_i^* \subseteq B_i$, $i=1, ..., n$, then $\Gamma^* \subseteq \Gamma$, where Γ and Γ^* are now the graphs associated respectively with rules $A_i \rightarrow B_i$ and $A_i \rightarrow B_i^*$. Indeed, $(x, z) \in \Gamma^*$ if and only if $\forall i,\ A_i(x) \le B_i^*(z)$. Since $\forall i,\ B_i^* \subseteq B_i$, it follows that $\forall i,\ A_i(x) \le B_i(z)$, i.e. $(x, z) \in \Gamma$.

The combination of P1 and P2 leads to:

P3: If $A_i \subseteq A_i^*$ and $B_i^* \subseteq B_i$, $i=1, ..., n$, then $\Gamma^* \subseteq \Gamma$, where Γ and Γ^* are now the graphs respectively associated with rules $A_i \to B_i$ and $A_i^* \to B_i^*$.

Thus, according to the above inclusion properties, it is possible to design a family of nested graphs simply by building collections of nested fuzzy subsets on X and Z. Indeed, denote by $\{A_i^\lambda, \lambda \in [0,1]\}$ a family of fuzzy subsets on X such that $A_i^{\lambda'} \subseteq A_i^\lambda$ if $\lambda \ge \lambda'$ and $\{B_i^\lambda, \lambda \in [0,1]\}$ a family of fuzzy subsets on Z such that $B_i^\lambda \subseteq B_i^{\lambda'}$ if $\lambda \ge \lambda'$. The family of rules $A_i^\lambda \to B_i^\lambda$, $\lambda \in [0,1]$, is such that $\Gamma^\lambda \subseteq \Gamma^{\lambda'}$ if $\lambda \ge \lambda'$. Actually, such a construction of nested graphs simply expresses that implicative graphs increase in the sense of inclusion when underlying constraints become more permissive. More permissive rules are obtained cither by restricting input conditions further, or by enlarging output fuzzy sets.

Using a convex linear combination enables the automatic construction of such a collection of nested fuzzy subsets ranging from the lower bound of the family to the upper one. Applying such a technique results in the following proposal:

$$A_i^\lambda = (1-\lambda)A_i^0 \oplus \lambda A_i^1,\ \lambda \in [0,1],\ i=1, .., n \qquad (6)$$

where A_i^0 and A_i^1, such that $A_i^0 \subseteq A_i^1$, are the lower and upper bounds of the family and $\oplus$ denotes the extended sum of fuzzy numbers. In the same way, nested output fuzzy subsets can be built according to:

$$B_i^\lambda = (1-\lambda)B_i^0 \oplus \lambda B_i^1,\ \lambda \in [0,1],\ i=1, .., n \qquad (7)$$

where B_i^0 and B_i^1, such that $B_i^1 \subseteq B_i^0$, are the upper and lower bounds of the family. It should be noted that the inclusion ordering of the B_i^λ using λ is the converse of the one of the A_i^λ, due to opposite behaviors with respect to graph inclusion.

Using the so-built fuzzy subset families (see figure 10) results in the following graph inclusions:

$$\Gamma^1 \subseteq \Gamma^\lambda \subseteq \Gamma^{\lambda'} \subseteq \Gamma^0,\ \lambda, \lambda' \in [0,1] \text{ and } \lambda \ge \lambda' \qquad (8)$$

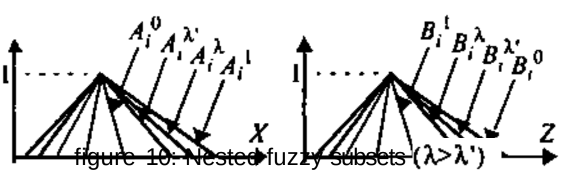


figure 10: Nested fuzzy subsets ($\lambda > \lambda'$)

One interesting point is that the 4-sided shape introduced in section 2 is shared by all nested graphs provided that the lower and upper graphs are themselves 4-sided areas. In other words, equalities (3) and (4) hold for any Γ^λ when they hold for Γ^0 and Γ^1 as expressed by property P4.

P4: If both (A_i^0, B_i^0) and (A_i^1, B_i^1), $i=1, .., n$, satisfy equations (3) and (4), then (3) and (4) are still valid for any $(A_i^\lambda, B_i^\lambda)$, $\lambda \in [0,1]$, $i=1, .., n$, when A_i^λ and B_i^λ are built according to (6) and (7).

An interpretation consists in viewing F as a fuzzy set of crisp graphs, that is as a level 2 fuzzy set, i.e. a fuzzy set of fuzzy sets. In this case, F is represented as:

$$F = \int_{\lambda \in [0,1]} \lambda / \Gamma^\lambda \qquad (9)$$

according to Zadch's notation where the integral sign stands for the union of the fuzzy singletons λ / Γ^λ. A single fuzzy gradual rule is pictured in figure 11.

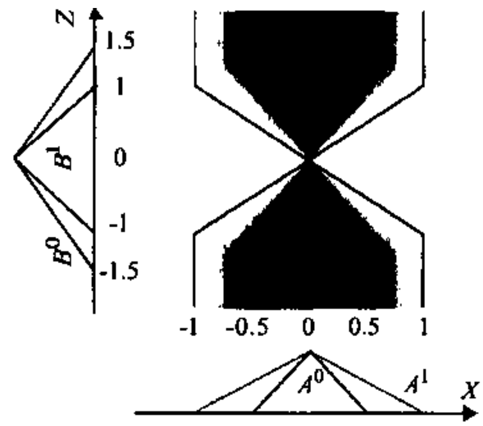


figure 11: A fuzzy gradual rule

Figure 12 plots the fuzzy graph obtained when the lower graph r^1 is precise and piecewise linear and the upper graph f° has the quadrangle-based shape of figure 2.

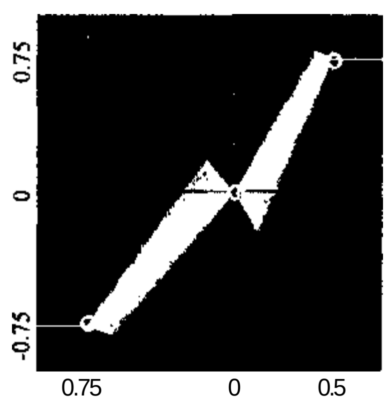


figure 12: A graph based on 3 fuzzy gradual rules

Using fuzzy gradual rules for dealing with the example of section 3 still results in perfect classification. Moreover, the robustness of the classification is improved as illustrated by figure 13, where histograms of the difference $N3 - N_5$ are plotted for class 3 time series (see figure 9). It is clear that the number of examples for which the final decision is brittle, i.e. for the small values of $N_3 - N_5$ which are accounted for in the two first bars, decreases when fuzzy gradual rules are used. These results have been obtained by using the same form of indices Nj given by equation (5) for crisp and fuzzy graphs except that $\Gamma_j(t_i, z_i) \in [0\ 1]$ in the fuzzy case. Further improvements could probably be obtained by defining more sophisticated indices in the fuzzy case.

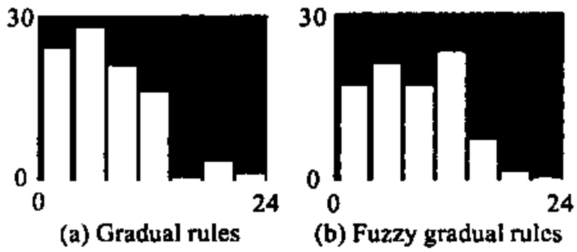


figure 13: Effect of fuzzy gradual rules on the discrimination

5 Conclusion

This paper has proposed a modelling framework which is faithful to the imprecision of available data. In the intervals between interpolation points where it is difficult to specify an analytical model, imprecision is captured by means of 4-sided areas. The application of the modelling methodology for classifying time series has exhibited interesting performance. Moreover, the discrimination power of the approach is improved by using fuzzy gradual rules as introduced in the previous section. Indeed, they enable us to distinguish between typical members of a class which remain in subareas with high membership degrees from borderline members which go through subareas with smaller membership degrees. Besides, one may probably take further advanvage of the easy interface with the user, provided by the use of gradual rule, for specifying queries in data mining applications (see [Keogh et al., 2002] for an example of such possible use). Lastly, further research should deal with multidimensional spaces where the language of gradual rules may prove useful in the description of imprecise graphs.

References

[Biswas *et al,* 1998] G. Biswas, J. Weinberg and D. Fisher. ITERATE: A conceptual clustering algorithm for data mining. *IEEE Trans, on Systems, Man and Cybernetics - Part C: Applications and reviews,* 28(2), 219-230, May 1998.

[Colomer *et al,* 2002] J. Colomer, J. Melendez, F.I. Gamero. Pattern recognition based on episodes and DTW. Application to diagnosis of a level control system. *Proceedings of the 16th Int. Workshop on Qualitative Reasoning (QR'02),* Barcelona, Spain, June 2002.

[Dubois and Prade, 1992] D. Dubois and H. Prade. Gradual inference rules in approximate reasoning. *Information Sciences,* 61(1,2), 103-122, April 1992.

[Dubois and Prade, 1994] D. Dubois and H. Prade. Basic issues on fuzzy rules and their application to fuzzy control. In *Fuzzy Logic and Fuzzy Control,* D. Driankov, P.W. Eklund, A.L. Ralescu Eds, Lecture Notes in Artificial Intelligence, 833, Springer-Verlag, 3-14, 1994.

[Dubois *et al,* 1997] D. Dubois, H. Prade and L. Ughetto. Checking the Coherence and Redundancy of Fuzzy Knowledge Bases. *IEEE Trans, on Fuzzy Systems,* 5(3), 398-417, August. 1997.

[Galichet *et al,* 2002] S. Galichet, D. Dubois and H. Prade. Imprecise specification of ill-known functions using gradual rules. *Proceedings of the 2nd Int. Workshop on Hybrid Methods for Adaptive Systems (EUNITE 2002),* pages 512-520, Albufeira, Portugal, September 2002.

[Kadous, 2002] M. W. Kadous. Temporal classification: extending the classification paradigm to multivariate time series. PhD thesis, The University of New South Wales, January 2002.

[Keogh and Kasetty, 2002] E. Keogh and S. Kasetty. On the need for time series data mining benchmarks: a survey and empirical demonstration. *Proceedings of the 8th ACM SIGKDD Int. Conf on Knowledge Discovery and Data Mining,* pages 102-111, Edmonton, Canada, July 2002.

[Kawaguchi and Miyakoshi, 1998] M.F. Kawaguchi and M. Miyakoshi. Fuzzy spline interpolation in sparse fuzzy rule bases. *Proceedings of the 5th International Conference on Soft Computing (IIZUKA '98),* pages 664-667, Iizuka, Japan, October 1998.

[Keogh *et al,* 2002] E. Keogh, H. Hochhciser and B. Shnciderman. An augmented visual query mechanism for finding patterns in time series data. *Proceedings of the 5th Int. Conf on Flexible Query Answering Systems (FQAS 2002),* pages 240-250, Copenhagen, Denmark, October 2002.

[Li, 2000] C. Li. A bayesian approach to temporal data clustering using the hidden Markov model methodology. PhD thesis, Vanderbilt University, December 2000.

[Lowen, 1990] R. Lowen. A fuzzy Lagrange interpolation theorem. *Fuzzy Sets and Systems,* 34, 33-38, 1990.

[Nanopoulos *et al,* 2001] A. Nanopoulos, R. Alcock and Y. Manolopoulos. Feature-based classification of time-series data. *International Journal of Computer Research,* 10(3), 2001.

[Sala and Albertos, 2001] A. Sala and P. Albertos. Inference error minimisation: fuzzy modelling of ambiguous functions. *Fuzzy Sets and Systems,* 121(1), 95-111, July 2001.

[UCI KDD Archive] The UCI KDD Archive (http://kdd.ics.uci.edu/), Information and Computer Science, University of California, Irvine.