Fuzzy gradual patterns: What fuzzy modality for what result?

Sarra Ayouni, Sadok Ben Yahia, Anne Laurent, Pascal Poncelet

HAL Id: lirmm-00798797
https://hal-lirmm.ccsd.cnrs.fr/lirmm-00798797

Submitted on 3 Nov 2019

Fuzzy Gradual Patterns:
What Fuzzy Modality For What Result?

Sarra AYOUNI*,**, Sadok BEN YAHIA*
*Faculty of Sciences of Tunis
Tunis, Tunisia
{sarra.ayouni,sadok.benyahia}@fst.rnu.tn

Anne LAURENT**, Pascal PONCELET**
**LIRMM - Univ. Montpellier 2/CNRS
Montpellier, France
{ayouni,laurent,poncelet}@lirmm.fr

Abstract

Gradual Patterns allow to express frequent patterns like "the older an employee, the higher his/her income". *After having been extensively used for fuzzy command (e.g.,* "the closer the wall, the harder the break"*), they are currently grasping the attention of the data mining community, where the focus is put on the scalability of the methods when addressing very large databases (e.g. biology). However, it is often the case that the information is not contained at the upper and lower boundaries (min/max), but rather hidden in-between. For instance, the most interesting rule could be* "the closer the age of an employee to 46, the higher his/her income". *In this paper, we introduce several methods for automatically defining fuzzy modalities and we compare the results obtained when considering the various solutions.*

1. Introduction

Gradual patterns are of the form *"the more/less A_1, ..., the more/less A_k"* [8], [12]. For instance, *"the older an employee, the higher his/her income"* refers to such a pattern. They have been recently reintroduced in the literature as one of the types of patterns that data mining techniques can be of help to discover [3], [6], [7], [20].

Standing within fuzzy settings, fuzzy gradual patterns are of the form *"the more/less A_1 is F_1 ..., the more/less A_k is F_k"* [8], [12], as for instance in *"the more an employee is almost 46 years old, the more his/her income is almost 37,000 euros per year"*, where *almost 46 years old* and *almost 37,000 euros* are fuzzy sets described by their respective membership functions. The fuzzy set theory, introduced by Zadeh [23], allows to deal with natural language and to imprecise concepts like *young* or *almost X years old*. A membership function μ of a fuzzy set A is an application from the domain D of the attribute being considered (*e.g.,* ages from 0 to 125) to $[0, 1]$. Considering a value $x \in D$, $\mu(x)$ ranges in $[0, 1]$ and indicates to which extent x belongs to the fuzzy set A.

Discovering fuzzy gradual patterns can be of benefit to many application areas. For instance, in biology, it is of

Table 1. Sample Data

Id	Age	Income (keuros)	Nb of Cars
1	32	30	1
2	43	40	2
3	48	37	2
4	46	39	3
5	57	34	3

paramount interest to retrieve the information that *"the closer the expression of gene i to .32, the closer the expression of gene j to $-.12$"*.

In this paper, we aim at building fuzzy modalities that seems to be the most relevant. Instead of building a complete partition, we rather focus on a fuzzy modality that is meant to gather the identity of the numerical attribute being considered. For doing this, our idea is to define fuzzy modalities that describe to which extent the value of an attribute is close to this attribute's median.

For instance, Table 1 displays sample data that describes people regarding their identifier, their age, their annual income, and the number of their cars. Unfortunately, it is not possible to retrieve gradual patterns from this table. However, if considering the fuzzy modalities *almost 46 years old* and *almost 37,000 euros* as sketched by Figure 1, the table can be transformed as showed by Table 2. Thus, the pattern *"the more an employee is almost 46 years old, the more his/her income is almost 37,000 euros per year"* appears to be true: the lines can be ordered considering the list of id $< 1, 5, 2, 3, 4 >$ and it can be seen that the pairs of values (degree of age almost 46, degree of income almost 37,000) appears to be ordered by a precedence order. Indeed, if we consider that a tuple t precedes another tuple t' (denoted as $t \triangleleft t'$) whenever all the components can be ranked in this order, then we have that tuples defined by projecting the database on (degree of age almost 46, degree of income almost 37,000) are ordered: $(0, 0) \triangleleft (0, .25) \triangleleft (.25, .25) \triangleleft (.5, .5) \triangleleft (1, 1)$.

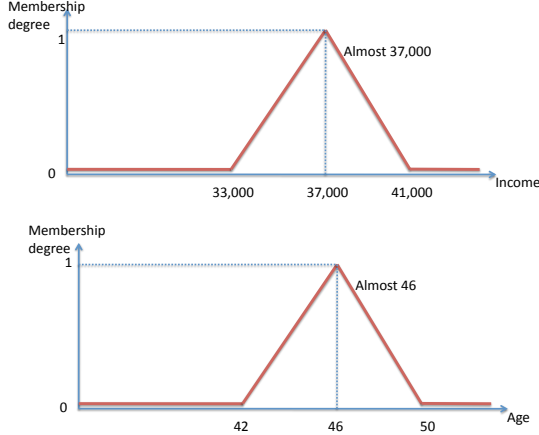Our objective is then to automatically retrieve fuzzy

Figure 1. Fuzzy Modalities over Age and Income attributes

Table 2. Transforming the Sample Data

| Id | Age | Income (keuros) | Nb of Cars |
	Almost 46	Almost 37	Almost 2
1	0	0	0.5
2	.25	.25	1
3	.5	.5	1
4	1	1	0.5
5	0	.25	0.5

modalities that allow to discover relevant fuzzy gradual patterns. We thus begin by recalling below some preliminaries about gradual patterns and automatic fuzzy partition building, before introducing our contribution for building fuzzy modalities in the framework of fuzzy gradual patterns. Experiments have been carried on real datasets and are reported in the penultimate section. The last section concludes this paper and sketches some issues for future work.

2. Related Work

In this section, we present an overview of approaches dealing with gradual patterns. We also present an overview of different strategies for building fuzzy membership functions.

2.1. Gradual Patterns

Gradual patterns are extracted from datasets described over m attributes $(X_1, \ldots, X_m)$ defined on numerical domains $\text{dom}(X_i)$. A dataset $\mathcal{D}$ is a set of rows (m-tuples) in $\text{dom}(X_1) \times \ldots, \times \text{dom}(X_m)$. In the case of fuzzy data, the X_i attributes are fuzzy linguistic variables associated to linguistic values defined through fuzzy sets. Let us consider an attribute corresponding to the speed of vehicles. In the

crisp case, it contains the numerical values of the measured speeds. Whereas, in the fuzzy case, it will be associated to linguistic variables, *e.g.,* "*slow*", "*normal*", and "*fast*". These linguistic variables or modalities are defined by their membership degrees (according to a membership function) indicating the extent to which their speeds belong to each modality. Extracting gradual patterns grasped the interest of many researchers [6], [7], [17], [18], [20] and various definitions were proposed to assess the support value of a gradual pattern.

A gradual pattern or gradual itemset is defined as a non-empty set of several gradual items. A gradual item is defined as a pair of an attribute and a variation $* \in \{\leq, \geq\}$. Let A be an attribute, then the gradual item $(A, \geq)$ means that the attribute A is increasing. It can be interpreted by "*the more A*". Here, the attribute A can be crisp or fuzzy. For instance (*Speed*, $\geq$) is a crisp gradual item, as well as (*Fast Speed*, $\geq$) which is a fuzzy gradual item.

A gradual rule, denoted by $M \Rightarrow M'$, is defined as a pair of gradual itemsets on which a causality relationship is imposed. Different measures and semantics have been proposed to extract and assess such kind of rules. Nevertheless, many semantics have been formalized to define the cases where such a rule can be said to hold or not. The authors in [18] propose a method to extract so called gradual dependencies denoted by $A \rightharpoonup_t B$. A linear regression analysis is performed on the contingency diagram depicted from the data set. The validity of the rule is assessed on the basis of the regression coefficients α, β of the line that approximates the points in the contingency diagram and the quality of the regression as given by the R^2 coefficient. A tendency rule contains one or more attributes in the condition part and only one in the conclusion part. When it contains many attributes in the condition part, the authors propose to use a logical conjunction modeled by means of a so-called *t-norm*. The gradual tendency is then understood as a covariance constraint between the aggregated membership degrees.

Another definition has been proposed in [3]. The semantics of a gradual dependence is quite different, since the authors only consider the fulfilment variation. The authors define a gradual dependence as being similar to a functional dependence. However, in a functional dependency, the equalities between the values are considered while in a gradual dependency the variations of degrees between two objects are of interest. According to [3], the gradual dependence $A \Rightarrow B$ holds in a database $\mathcal{D}$ if $\forall$ $o=(x, y)$ and $o'=(x', y') \in \mathcal{D}$, $A(x) < A(x')$ implies $B(y) < B(y')$. It should be noted that this definition does not take into account any comparison between data rows.
A new definition of gradual dependencies was proposed in [21] using fuzzy association rules. The authors take into account the variation strength in the degree of fulfilment of

an imprecise property by different objects. Hence, a gradual dependency holds in a database $\mathcal{D}$ if $\forall\ o=(x, y)$ and $o'=(x', y') \in \mathcal{D}$, $v_{*1}(A(x), A(x'))$ implies $v_{*2}(B(y), B(y'))$, where v_* is a variation degree of an attribute between two different rows. In both propositions [3] and [21], the authors propose to build a modified data set $\mathcal{D}'$ that contains as many rows as there are pairs of distinct objects in the initial data set $\mathcal{D}$ in order to identify such gradual dependencies.

Another definition was proposed in [6]. As pointed out in the introduction, the support of a gradual itemset $A_1^{*_1}, \ldots, A_p^{*_p}$ is defined as the maximal number of rows $\{r_1, \ldots, r_l\}$ for which there exists a permutation π such that $\forall\ j \in [1,\ l-1]$, $\forall\ k \in [1,\ p]$, it holds $A_k(r_{\pi_j})\ *_k\ A_k(r_{\pi_{j+1}})$. More formally, denoting $\mathcal{L}$ the set of all such sets of rows the support of a gradual itemset is defined as follows. Let $s = A_1^{*_1}, \ldots, A_p^{*_p}$ be a gradual itemset, we have:

$$supp(s) = \frac{\max_{L_i \in \mathcal{L}} |L_i|}{|\mathcal{D}|}$$

The authors propose a heuristic to compute this support for gradual itemsets, in a levelwise process that considers itemsets of increasing lengths. The authors in [7] propose the GRITE method [1] which is an efficient method based on the precedence graph and binary matrices for data representation.

In [20], the authors proposed a novel definition of the support of gradual patterns based on rank correlation (concordant and discordant pairs of rows from the data). The authors use the same algorithm and data binary representation as cited above.

Another type of fuzzy gradual rules was proposed in [5]. The authors propose to mine gradual rules enriched with a clause introduced by the expression "all the more": such rules of the form *"the more X is A, the more Y is B, all the more Z is C"* offer additional precisions on the relation between the attributes. For instance the pattern *"the closer the wall, the harder the brakes are applied, all the more the higher the speed"* means that the pattern *"the closer the wall, the harder the brakes are applied"* is more true when the speed is high than when the speed is not high. It should be noted that many interpretations and computations could be proposed in this framework.

We focus below on the methods that have been proposed for building membership functions.

2.2. Building Fuzzy Modalities

In this respect, most of the methods build either a triangular, trapezoidal or gaussian function. If considering the problem of building a triangular membership function, then the problem will be to delimit the support and kernel of the fuzzy set. Let us consider that the support is defined
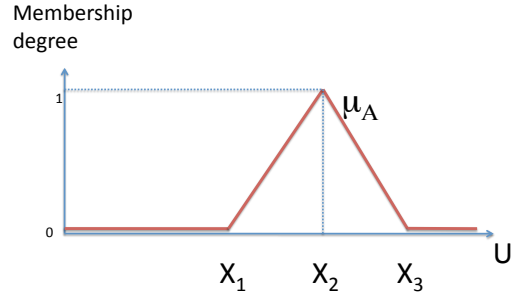
Figure 2. An example of a Triangular Membership Function

in the interval $[x_1 x_3]$ and the kernel is $\{x_2\}$ (see Figure 2). The function is defined by:

$$\mu_A(x) = \begin{cases} \frac{x-x_1}{x_2-x_1} & if\ x_1 \leq x < x_2 \\ \frac{x-x_3}{x_2-x_3} & if\ x_2 \leq x < x_3 \\ 0 & otherwise \end{cases}$$

The values have to satisfy the following condition:

$$x_1 < x_2 < x_3$$

If considering the problem of building a trapezoidal membership function, then the support is defined in the interval $[x_1, x_4]$ and the kernel is $[x_2, x_3]$ (see Figure 3). The function is defined by:

$$\mu_A(x) = \begin{cases} \frac{x-x_1}{x_2-x_1} & if\ x_1 \leq x < x_2 \\ 1 & if\ x_2 \leq x < x_3 \\ \frac{x-x_4}{x_3-x_4} & if\ x_3 \leq x < x_4 \\ 0 & otherwise \end{cases}$$

The values have to satisfy the following condition:

$$x_1 < x_2 < x_3 < x_4$$

If considering the problem of building a gaussian membership function, then we consider a center x_2 and parameters x_2 and x_3 (see Figure 4). The function is defined by:

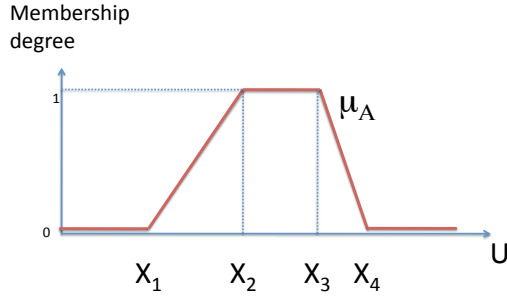$$\mu_A(x) = e^{\left(\frac{x-x_2}{w}\right)^2}$$

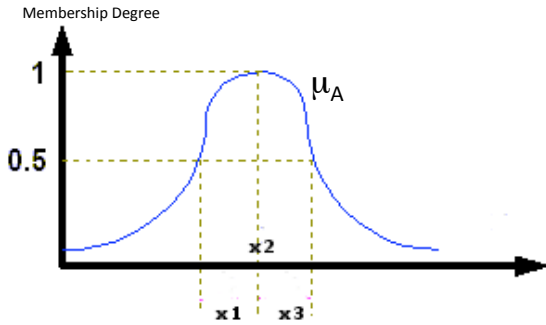Figure 3. An example of a Trapezoidal Membership Function

Figure 4. An example of a Gaussian Membership Function

Many methods were proposed in the literature to build fuzzy membership functions. These methods have been classified into four categories in [1]: automatic methods, statistical and probabilistic methods, psychometric methods and geometric methods, as reviewed below.

2.2.1. Automatic Methods. Automatic methods to construct membership functions are based on a two-step process that first defines a primary membership function (*e.g.,* by random) that is then adjusted in the second step.

The second step relies on an optimization performed by methods such as neural networks or evolutionary algorithms. Evolutionary algorithms are inspired by biological evolution (reproduction, selection, mutation) [14], [16]. They rely on a population that is enhanced step by step by combining the elements and by selecting the best ones. A fitness function

must be defined in order to rank the elements regarding their performance. For instance, in the context of data mining, the aim may be to increase the number of large itemsets and/or rules discovered or their confidence [19], [2].

Regarding the problem of building membership functions, the elements of the population (*i.e.,* the chromosomes) are randomly generated and each chromosome represents the parameters describing a membership function.

After the encoding step, these chromosomes undergo through a selection (individual solutions are selected through a fitness-based process) and reproduction steps. In this later step, a second population is generated from those selected through genetic operators (crossover, and mutation). This generational process is repeated until a termination condition has been reached (*e.g.,* a number of runs, or a indicator that the membership function is satisfactory).

2.2.2. Clustering Methods. When using clustering, the membership functions are generated either during the process of the clustering [4], or on the basis of certain parameters diverted from the obtained clusters [10]. In these methods, membership functions can represent the adequacy rate of each element to the cluster to which it is attached. It can also be the distance between every element and the center of gravity of its cluster.

2.2.3. Statistical and Psychometric Methods. Statistic methods rely on the study of the distribution of the values to build the membership function. For instance, histograms are used and a smooth function is designed on them. "*Yes-No*" methods consist in asking every member of the population a question having yes or no answer. The membership degree of an element to a fuzzy set is built as being equal to the number of the "yes" answers [11], [13], [15]. In a statistical method based on set estimation, the degrees of membership functions are directly given by individuals. The question "what is the degree of membership of an element x in the fuzzy set A?" is repeated several times in different moments to one or several person(s). A variable is used to register the obtained answers. The membership function is then defined by a conditional distribution function $f(y/x)$ where y is a variable registering the answers [22].

3. Discovering Fuzzy Gradual Patterns

We recall that our goal is to compare the gradual patterns obtained on numerical databases when applying the existing approaches on fuzzified data. We aim at discovering patterns including linguistic terms that describe values that are in-between the two extremes of the minimum and maximum values of the attributes, e.g., patterns like *"the closer the age of an employee to 46, the closer his/her income to 37,000 euros per year"*.

3.1. Building the Fuzzy Modalities

We study here some means to define the fuzzy set. Several possibilities are explored.

It is important to note that it would be hardly possible to ask an expert to design a fuzzy partition or fuzzy modality on all the attributes of a database, for instance a genomic database containing the expression of thousands of genes. The expert would then been asked to define a single fuzzy description that will be applied on the whole dataset, which we argue to be non relevant.

Here the aim is rather to obtain automatically the *best* fuzzy modality. However, it is very difficult to define what would be this best fuzzy modality. We argue here that this modality must allow to discover some interesting patterns that would not have been retrieved by existing methods, either because no partition was defined, as in the crisp approaches [6], or because the fuzzy partition being considered is pre-defined and not automatically built [3].

Moreover, this fuzzy description must cover the whole universe and must be interpretable as "almost v" in order to simplify the reading of the fuzzy gradual patterns to be extracted.

Finally, our method must remain scalable, as this is one of the main challenges of fuzzy data mining.

We thus focus on a method which is based on the definition of a function describing a smooth transition between all the values. As the values are split around the median, we consider the median as the central point of our fuzzy description and we then describe all the values depending on the extent to which they are similar to the median. For this purpose, several possibilities are explored. We may consider then several ways to build membership functions, for instance gaussian membership function or triangular membership function where the center is the median and the width is the whole universe in cover all the values.

However, there is no difference in the ranking of the values when we consider triangular or gaussian functions if considering the same value to the center as they are piecewise monotonic.

It should be noted that we aim at retrieving the most interesting fuzzy modality for *every* attribute. It can thus be the case that two attributes describe the same type of data (*e.g.,* expression of gene), but that the fuzzy modality is not the same one for both of them. This allows to distinguish our method to classical ones and to be as close as possible to the dataset.

Some other methods could be considered to build the membership functions, as for instance genetic programming. However, these methods are very time consuming and could not scale on databases containing thousands of attributes.

The median may be replaced by the *average* of the variable or by the most frequent value (although this is usually hardly reasonable on quantitative attributes), and we may consider the standard deviation for defining the width of the support, instead of covering all the universe. However, this would not change much the algorithms described below.

3.2. Algorithms

Our method stands within a pre-treatment step, that can be applied before all existing algorithms for gradual pattern discovery. The runtimes are thus not an indicator of the quality of the method building the fuzzy modalities but they must remain as short as possible. This pre-treatment has been implemented using the R software[2], which makes it possible to efficiently manage large databases within reasonable runtimes.

The algorithm 1 proceeds in two steps as described below. The first one (*i.e.*; Algorithm 2) corresponds to fuzzy modalities building and Algorithm 3 refers to classical algorithm to extract gradual patterns, as that given in [7].

The $BuildFuzzMod$ function as described in Algorithm 2 computes the key parameters of the fuzzy modality of each attributes of the original dataset. After that it returns the membership degree of each value of the attribute to the corresponding fuzzy modality that has been found.

Algorithme 1 Mining Fuzzy Gradual Patterns

---

1 `INPUT:`D `Input Database,` ς `MinSup`
2 `OUTPUT:` F `Set of frequent gradual`
   `patterns`
3 $D' \leftarrow BuildFuzzMod(D)$
4 $F \leftarrow$ `MineGradualPatterns(D',`ς`)`
5 `Return` F

---

4. Experimental Results

Our approach has been tested on several databases. One of them contains data on NBA players[3] containing 17 dimensions (player_regular_season). A second one describes the behavior of 500 genes for a biological study on breast cancer, while a third one describes data about wine [9] (13 columns) where the class attribute has been removed.

All these three datasets are pre-treated in order to replace every attribute by a fuzzy set describing to which extent the initial value is close to the median. For instance, Table 3 displays the new attributes for database $wine$. The value being considered is the median.

---

2. R is a free software environment for statistical computing and graphic and it can be downloaded on http://www.r-project.org

3. Statistics about NBA players are available at http://www.databasebasketball.com/stats_download.htm

Table 3. Wine Gradual Items Before and After Pretreatment

Attribute	Classical Description	Fuzzy Description
Malic Acid	The More/Less the Malic Acid	The more/less the Malic Acid is *almost* 13.05
Ash	The More/Less the Ash	The more/less the Ash is *almost* 1.865
Alcalinity of ash	The More/Less the Alcalinity of ash	The more/less the Alcalinity of ash is *almost* 2.36
Magnesium	The More/Less the Magnesium	The more/less the Magnesium is *almost* 19.5
Total phenols	The More/Less the Total phenols	The more/less the Total phenols is *almost* 98
Flavanoids	The More/Less the Flavanoids	The more/less the Flavanoids is *almost* 2.355
Nonflavanoid phenols	The More/Less the Nonflavanoid phenols	The more/less the Nonflavanoid phenols is *almost* 0.34
Proanthocyanins	The More/Less the Proanthocyanins	The more/less the Proanthocyanins is *almost* 1.555
Color intensity	The More/Less the Color intensity	The more/less the Color intensity is *almost* 4.69
Hue	The More/Less the Hue	The more/less the Hue is *almost* 0.965
OD280/OD315 of diluted wines	The More/Less the OD280/OD315 of diluted wines	The more/less the OD280/OD315 of diluted wines is *almost* 2.78
Proline	The More/Less the Proline	The more/less the Proline is *almost* 673.5

Algorithme 2 BuildFuzzMod - Building Fuzzy Modalities

1. INPUT: D Input Database
2. OUTPUT: D' Transformed Database
3. **foreach** Attribute A **in** D **do**
4.   TmpMin $\leftarrow$ Min($D.A$)
5.   TmpMax $\leftarrow$ Max($D.A$)
6.   TmpCenter $\leftarrow$ Median($D.A$)
7.   // Build the triangular membership function MFunc
8.   MFunc $\leftarrow$ Triang(TmpMin,TmpCenter,TmpMax)
9. **endforeach**
10. **foreach** Row r **in** D **do**
11.   **foreach** Attribute A **in** D **do**
12. $D'[A,r] \leftarrow$ MFunc(D[A,r])
13.   **endforeach**
14. **endforeach**
15. Return D'

Algorithme 3 MineGradualPatterns - Mining Gradual Patterns

1. INPUT: D Input Database, ς MinSup
2. OUTPUT: F Set of frequent gradual patterns
3. $F \leftarrow$ Pairs of Attributes Such that Support $\geq \varsigma$
4. **foreach** Size k **in** $[3,...]$ **do**
5.   Build candidates from Frequent of Size $k-1$
6. $F \leftarrow F \bigcup$ {Gradual Patterns containing k attributes such that Support $\geq \varsigma$}
7. **endforeach**
8. Return F

Table 4. Number of Patterns Extracted w.r.t. the $MinSup$ Value

	wine_crisp	wine_fuzzy	nba_crisp	nba_fuzzy
1	23	38,259	0	275
0.8	149	797,148	0	19,040

Table 4 displays the number of patterns discovered with and without considering a fuzzy modality for two different values of minimum support. In fact, crisp gradual patterns are of the form *"the more/less A_1, ..., the more/less A_k"*, while fuzzy gradual patterns are of the form *"the more/less A_1 is almost X, ..., the more/less A_k is almost Y"*. In order to discover crisp patterns, the GRITE method [7] has been used. It can be seen that our method allows to retrieve some patterns that could not be extracted before.

Table 3 reports some of the patterns discovered using our method. For the wine database, attributes *Total phenols* and *Flavanoids* are not linked with the crisp approach, while building a fuzzy modality allows to discover that they are correlated when considering *The more/less the Total phenols is* almost 98 and *The more/less the Flavanoids is* almost 2.355. Based on this correlation, patterns containing 9 attributes have been built.

For the NBA database, no pattern was built for a minimum support of 100% while 275 patterns were built considering the fuzzy modalities. For instance, attributes 1 and 12 have been retrieved to be correlated: *the more* Games Played *is almost* 122*, the less* Field Goals Attempted *is almost* 626. To sum up, we can conclude that our approach allows to discover potentially useful hidden knowledge in large datasets that crisp approach can not discover.

5. Conclusion

In this paper, we propose a method for automatically building fuzzy modalities from numerical datasets in order to build fuzzy gradual patterns. Many perspectives are linked

to this work. Other methods and statistics may be used. The goal may also be changed in order, for instance, to build the fuzzy membership function that allows to retrieve the highest number of rules, or the most discriminant rules for classifying data (*e.g.,* wine). Moreover, each attribute may be divided into several fuzzy subsets, for building a complete fuzzy partition and exploiting it. Genetic programming may also be used to automatically build the partition, as done for fuzzy association rules. Finally, user preferences may be taken into account, together with closed patterns, so as retrieving as few patterns as possible, in order to avoid drowning the user in the resulting patterns while retrieving the most relevant ones.

References

[1] N. Aladenise and B. Bouchon-Meunier. Acquisition de connaissances imparfaites : mise en évidence d'une fonction d'appartenance. *Revue Internationale de Systémique*, 11(1):109–127, 1997.

[2] J. Alcalá-Fdez, R. Alcalá, M. Gacto, and F. Herrera. Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms. *Fuzzy Sets Syst.*, 160(7):905–921, 2009.

[3] F. Berzal, J.-C. Cubero, D. Sanchez, M.-A. Vila, and J. M. Serrano. An alternative approach to discover gradual dependencies. *Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15(5):559–570, 2007.

[4] J. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[5] B. Bouchon-Meunier, A. Laurent, M.-J. Lesot, and M. Rifqi. Strengthening fuzzy gradual rules through "all the more" clauses. In *IEEE World Congress on Computational Intelligence - Fuzz'IEEE*, 2010.

[6] L. Di Jorio, A. Laurent, and M. Teisseire. Fast extraction of gradual association rules: A heuristic based method. In *Proc. of CSTST'08*, 2008.

[7] L. Di Jorio, A. Laurent, and M. Teisseire. Mining frequent gradual itemsets from large databases. In *Proc. of IDA'09*, 2009.

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

[9] A. Frank and A. Asuncion. UCI machine learning repository, 2010.

[10] A. Fu, M. Wong, S. Sze, and W. Wong. Finding fuzzy sets for the mining of fuzzy association rules for numerical attributes. In *Proceedings of the first International Symposium of Intelligent Data Engeneering and Learning (IDEAL'98)*, pages 263–268, October 1998.

[11] B. R. Gaines. Fuzzy and probability uncertainty logics. In *Information and Control*, volume 38, pages 297–323. 1987.

[12] S. Galichet, D. Dubois, and H. Prade. Imprecise specification of ill-known functions using gradual rules. *Int. Journal of Approximate Reasoning*, 35(3):205–222, 2004.

[13] R. Giles. The concept of grade of membership. *Fuzzy Sets Syst.*, 25(3):297–323, 1988.

[14] D. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Professional, Upper Saddle River,NJ, USA, 1989.

[15] E. Hisdal. Are grades of membership probabilities? *Fuzzy Sets Syst.*, 25(3):325–348, 1988.

[16] J. Holland. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press, Cambridge Mass., 1st MIT press ed. edition, 1992.

[17] E. Hüllermeier. Implication-based fuzzy association rules. In *Proc. of PKDD'01*, pages 241–252, 2001.

[18] E. Hüllermeier. Association rules for expressing gradual dependencies. In *Proc. of PKDD'02*, pages 200–211, 2002.

[19] M. Kaya and R. Alhajj. Utilizing genetic algorithms to optimize membership functions for fuzzy weighted association rules mining. *Applied Intelligence*, 24(1):7–15, 2006.

[20] A. Laurent, M.-J. Lesot, and M. Rifqi. Graank: Exploiting rank correlations for extracting gradual dependencies. In *Proc. of FQAS'09*, 2009.

[21] C. Molina, J. Serrano, D. Snchez, and M. Vila. Measuring variation strength in gradual dependencies. In *Proceedings of the International Conference EUSFLAT'2007, Ostrava, Czech Republic*, pages 337–344, 2007.

[22] P. Wang. From the fuzzy statistics to the falling random subsets. *Advances in Fuzzy Sets, Possibility Theory and Applications*, pages 81–96, 1983.

[23] L. Zadeh. Fuzzy sets. *Information and Control Journal*, 8:338–353, 1965.