Towards Fuzzy-OLAP Mining

A. Laurent, B. Bouchon-Meunier, and A. Doucet

Université Pierre et Marie Curie
LIP6 - Dept. of Artificial Intelligence
8, rue du capitaine Scott
F-75015 PARIS - FRANCE
Anne.Laurent@lip6.fr

Abstract. Multidimensional databases and OLAP (*On Line Analytical Processing*) tools provide an efficient framework for data mining and have led to the so-called *OLAP Mining* architecture ([10]). Besides, data from real world are often imperfect, either because they are uncertain, or because they are imprecise. Moreover, the use of fuzzy set theory in data mining systems enhances the understandability of the discovered knowledge when considering numerical attributes and leads to more generalizable rules. This is the reason why we have proposed an approach to perform OLAP-based mining using multidimensional databases and fuzzy decision trees ([17]). In recent works, an extension of multidimensional databases has been defined in order to handle imperfect information and flexible multidimensional queries ([15]). In this paper, we show that this model is well adapted for data mining. A general architecture is proposed, which uses fuzzy multidimensional databases as a support for knowledge discovery.

1 Introduction

Learning systems provide efficient data mining tools when they are integrated with multidimensional databases. OLAP tools and multidimensional databases provide indeed a framework for hierarchies and enhance the calculus of aggregates.

Besides, the introduction of fuzziness in the data mining process is very interesting to deal with numerical attributes. In this framework, fuzzy set theory provides tools to manipulate data with words instead of using numerical descriptions, leading to more understandable knowledge. For instance, a rule like "*Most* sales in *eastern* cities are *high*" sounds very natural for the user, whereas rules like "The average of sales is 88,200 units" contain less information for analysts.

However, many fuzzy data mining systems extract knowledge from flat files. They do not take advantage of the performances of database management systems, and even less of the benefits from the multidimensional model.

Thus, we propose to integrate multidimensional databases with fuzzy learning systems. For this purpose, the multidimensional model has been extended to the management of imperfect data and flexible queries in [15] for the definition of fuzzy multidimensional databases. In this paper, we show that this

model provides the needed operators for its integration in the OLAP Mining architecture.

This solution has been implemented and tested using the OLAP software from `Oracle`: `Oracle Express Server`. We strongly insist on the fact that this model generalizes the classical one, and provides a framework to handle all classical crisp cases, since fuzzy set theory provides means to handle imperfect *and* classical data.

In the sequel, Section 2 briefly describes what fuzzy knowledge is. Section 3 presents existing work concerning the integration of multidimensional databases and data mining systems and existing work on fuzzy learning. Then Section 4 briefly introduces the model we propose for fuzzy multidimensional databases, and Section 5 details the architecture which integrates these fuzzy multidimensional databases with data mining systems. Section 6 details the data mining systems integrated in the architecture, and Section 7 presents the implementation of our solution. Finally, Section 8 concludes and presents further work.

2 Fuzzy Data and Flexible Queries

Fuzzy set theory provides a framework to handle data from real world. Data are indeed often imperfectly defined. They may be imprecise ("sales are *bad*") and/or uncertain ("sales *may* have been bad"). In the case of numerical attributes, we argue that it is not suitable to represent concepts from natural language such as *bad sales* by crisp partitions. It would lead to arbitrary classes that do not reflect the reality. For instance, there is no reason to consider a classical partition of the ages into crisp categories, since the transition areas between characterizations are obviously fuzzy.

Moreover, even in the cases where data are well defined, it may be interesting to apply flexible queries, for instance to select *middle* sales.

Finally, the possibility to represent fuzzy hierarchies is interesting. For instance, it is useful to represent a geographical attribute with some levels of abstraction, where cities may belong gradually to regions. This is very useful to prevent analysis from being led by administrative regions, instead of real geographical distribution. Two cities may be close, in the eastern part of the considered geographical area, even if they do not belong to the same administrative district. Furthermore, there is a need for values to belong to more than one concept. For instance, a 31-year-old people may be considered as *middle aged* and *young*, whereas a crisp partition where *young* ranges from 0 to 30 will consider him as *not young*.

Fuzzy set theory provides a theoretical framework to manipulate this kind of data in a rigorous way. In this theory, an element can belong gradually to a set. More formally, given a reference set X, a fuzzy set A from X is defined by a membership function μ_A that associates each element x from X with the degree $\mu_A(x)$, ranging from 0 to 1, which indicates at which extent the element x belongs to A. This theory enables also the representation of classical crisp data. Fig. 1 shows some examples of all types of values that may be represented.

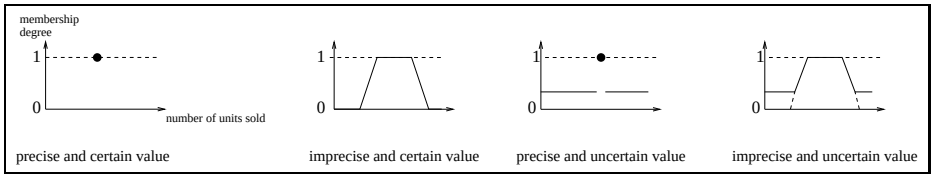

Fig. 1. Representation of data in fuzzy set theory

Flexible queries are evaluated by using measures of satisfiability ([6]), which enable us to compute the degree to which a (fuzzy or non fuzzy) value satisfies a fuzzy criterion. An example of such a measure is given by equation (1).

$$\mathcal{C}(A,B) = \frac{M(A \wedge B)}{M(A)} \text{ where } M(A) = \begin{cases} \int_{\Omega} \mu_A(x)dx \text{ if } \Omega \text{ is continuous} \\ \sum\limits_{x \in \Omega} \mu_A(x) \quad \text{otherwise} \end{cases} \quad (1)$$

This measure is considered in our model. It satisfies the following property in the case of a precise value $\{\dot{b}\}$: if $B = \{\dot{b}\}$ then $\mathcal{C}(A,B) = \mu_A(\dot{b})$

Fuzzy set theory also enables us to represent fuzzy quantifiers, such as *a few, most*, by membership functions (as illustrated by Fig. 2). *Absolute* quantifiers like *at least 2* are distinguished from *relative* ones (*most* for instance).

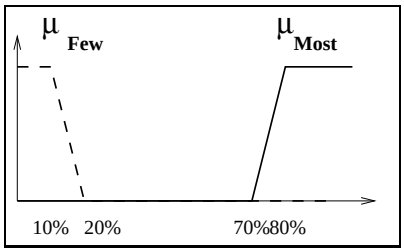

Fig. 2. Fuzzy Quantifiers

The management of fuzzy data and flexible queries has been studied in fuzzy databases ([2]). However, no study has been led on fuzzy multidimensional models. We propose thus a model where fuzziness is integrated in order to store fuzzy data and to apply fuzzy queries on cubes. The needed definitions and operations are detailed in Section 4. Then Section 5 shows that this model is suitable for knowledge discovery. Related work concerning the existing architectures for knowledge discovery using multidimensional databases or fuzzy systems are described in the following Section 3.

3 Related Work

This section presents existing work concerning OLAP Mining and fuzzy learning.

3.1 OLAP Mining

The amount of available data is growing dramatically. Large amounts of data coming from miscellaneous sources are mainly managed by data warehouses. Data warehouses provide solutions to extract massive amounts of data from heterogeneous data sources, to clean and transform them in order to load them in large repositories that are constituted by a set of materialized views. Data marts are built from data warehouses to address a particular set of questions for specific users. They contain less information and are devoted to a particular analysis task.

In this framework, OLAP applications have emerged in order to provide tools to manage data for analysis. The term OLAP, standing for *On-Line Analytical Processing*, and multidimensional databases have been proposed in 1993 ([7]). They address a category of applications devoted to data analysis.

Architectures have been proposed in order to perform data analysis on multidimensional data. The term *OLAP Mining* has been proposed by J. Han in [10] in order to address systems that integrate OLAP tools and data mining systems. Several data mining components have been integrated in this architecture ([11]), for instance association rules ([14]) and decision trees. These studies include works on multiple-level discovery of relevant knowledge ([12]).

3.2 Fuzzy Learning Systems

Many fuzzy systems have been proposed in order to discover fuzzy knowledge from data. In this paper, we focus on fuzzy decision trees and fuzzy summaries. However, many other systems exist, for instance fuzzy clustering.

Fuzzy Decision Trees

Decision trees are well-known tools to represent knowledge. Several inductive learning methods exist to construct a decision tree from a training set of data [19]. Edges of a tree can also be labeled by fuzzy values. Such a kind of values leads to the generalization of decision trees to fuzzy decision trees [5] that handle fuzzy values either during their construction or when classifying new cases. The use of fuzzy set theory enhances the understandability of decision trees when considering numerical attributes.

A decision tree is constructed from the root to the leaves, by successive partitioning of the training set into subsets. The construction process can be split in 3 fundamentals steps. An attribute is selected thanks to a measure of discrimination H (step 1) that orders the attributes according to their accuracy with regard to the class. The partitioning is done by means of a splitting strategy P (step 2). A stopping criterion T enables us to stop splitting a set and to construct a leaf in the tree (step 3). To build fuzzy decision trees, the commonly used measure is the star-entropy measure defined as:

$$H_S^*(c|A_j) = -\sum_{l=1}^{L} P^*(v_l) \sum_{k=1}^{K} P^*(c_k|v_l) \log(P^*(c_k|v_l)) \qquad (2)$$

where $v_1, ..., v_L$ are the L values for attribute A_j and $c_1, ..., c_K$ are the K values for the class c. This measure is obtained from the Shannon measure of entropy by introducing Zadeh's probability measure P^* of fuzzy events.

Fuzzy Summaries

Fuzzy summaries have been studied since the early 1980s ([3, 8, 9, 13]). They are very interesting since they describe data with words from natural language: *"Most sales are bad"*. They use fuzzy quantifiers, as described above, in section 2. For instance, given a quantifier Q and some objects y_i $(i = 1, \ldots, n)$ to describe, fuzzy summaries are then like: "Q y_i *are* $S : \tau$" or "Q B y_i *are* $S : \tau$" where B is a condition on objects, S is the summarizer, and τ is the degree of truth of the summary. τ is computed as follows:

$$\tau = \mu_Q\left(\frac{1}{n}\sum_{i=1}^{n}\mu_S(y_i)\right) \quad \text{or} \quad \tau = \mu_Q\left(\frac{1}{n}\sum_{i=1}^{n}\top(\mu_S(y_i), \mu_B(y_i))\right)$$

where μ_Q is the membership function associated with Q, μ_S the membership function associated with S and $\top$ is a particular aggregation function (a triangular norm), for instance min.

As this section shows, OLAP Mining tools are very interesting for knowledge discovery. Besides, fuzzy learning provides tools to handle imperfect data and to extract more understandable knowledge from databases. However, no system combines OLAP Mining and fuzzy data mining. We propose thus to integrate multidimensional databases and fuzzy learning systems. For this purpose, we have introduced a model for fuzzy multidimensional databases, presented below.

4 Fuzzy Multidimensional Databases

Many models have been proposed for multidimensional databases ([20]). Roughly speaking, a multidimensional database is a set of *hypercubes* (hereafter *cubes*), defined on dimensions, which may be organized hierarchically. One dimension of particular interest is chosen as the *measure* whose values are stored in the cube cells. Operations are defined to visualize and manipulate cubes (*e.g.* rotation, selection by slice and dice, roll-up, drill-down).

In a previous paper ([15]), we have extended this model in order to handle imperfect data and flexible operations. We describe briefly this model below.

Entities Data in the cubes may be imprecise and/or uncertain. For this reason, we call *element* a pair (v, d) where v is a label standing for a fuzzy set (or a precise value) and d ($d \in [0, 1]$) is the degree of confidence in this value. We call domain of a dimension a finite set of such elements.

When the dimension is the measure, each cell is described by its position on each other dimension. $\overrightarrow{x}$ stands for such a cell, and we note $(v(\overrightarrow{x}), d(\overrightarrow{x}))$ the corresponding entity which constitutes the cell element. On dimensions, we note

d_i an entity, and we note $(v(d_i), d(d_i))$ the pair representing the corresponding element.

Each cell $\overrightarrow{x}$ of a fuzzy cube is associated with a degree $\mu(\overrightarrow{x})$ corresponding to the degree to which it belongs to the cube.

For instance, a fuzzy cube C is defined as a relation $C : PRODUCT \times MONTH \times DISTRICT \rightarrow SALES \times [0,1]$. Fig. 3 shows an example of such a fuzzy cube.

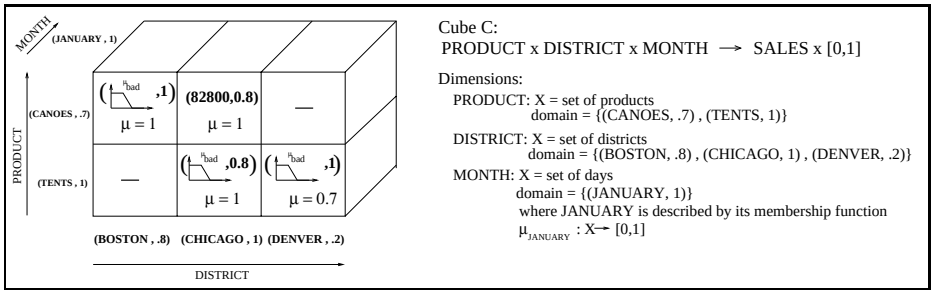

Fig. 3. A fuzzy cube

Hierarchies Hierarchies may be defined on dimensions, in order to represent data at different levels of granularity (see Fig. 4). These hierarchies may be defined using either fuzzy relations or fuzzy partitions ordered by a relation $\prec$.

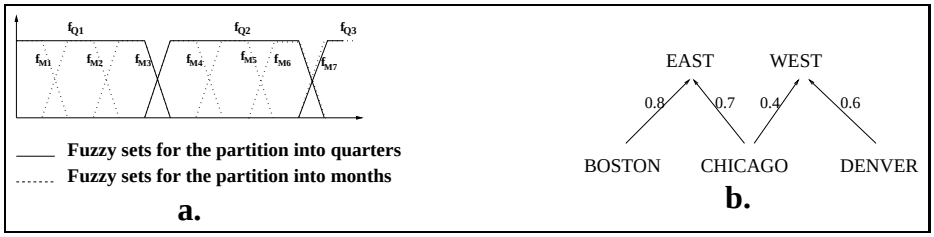

Fig. 4. Fuzzy hierarchies

Operations We define operations on such fuzzy cubes. Some of these operations are briefly presented here, but they are not detailed. See [15] for further information.

Selection on cell values (*dice*) This operation is defined in order to select cells from the cube matching some criterion that may be imprecise. For instance, *middle* sales are selected from the previous example cube.

This operation results in a new cube where the μ degrees have been modified in order to take into account the extent to which the entity matches the selection criterion.

Selection on dimension values (*slice*) Elements on dimensions belong gradually to the cube, depending on the $d(d_i)$ values. This operation modifies these degrees depending on how the element on the dimension matches the selection criterion. For instance, all slices corresponding to *eastern* districts are selected from the cube C.

Aggregation We call *aggregation* the operation consisting in summarizing a cube C by a value, fuzzy or not. For instance, we consider two kinds of counting. The first one sums all μ degrees from the cells, while the second one counts all cells having a μ degree greater than a given threshold t. For both counting methods, a cell is taken into account only if all corresponding slices belong to the cube with a degree greater than a given value χ:

$$\sum_{\overrightarrow{x}} \mu_C(\overrightarrow{x}) \times \prod_{i=1}^{k} \delta_\chi(d(d_i)) \qquad \text{or} \qquad \sum_{\overrightarrow{x}} \delta_t(\mu_C(\overrightarrow{x})) \times \prod_{i=1}^{k} \delta_\chi(d(d_i))$$

$$\text{where } \delta_\chi(x) = \begin{cases} 1 \text{ if } x \geq \chi \\ 0 \text{ otherwise} \end{cases}$$

Roll-Up Rolling up a cube consists in considering it at a higher level of granularity. For instance, the *sales* cube can be rolled up on the geographical dimension from the district level to the region one.

5 Fuzzy Data Mining from Fuzzy Multidimensional Databases: Architecture

The general architecture we propose is described in Fig. 5. It integrates fuzzy multidimensional databases, a fuzzy OLAP query interface, and learning systems. The requirements for such an architecture are the following ones:

– selection on cell values and on dimension values (slice and dice) using fuzzy criteria,

– representation of results of queries in order to apply successive operations to hypercubes,

– representation of imperfect data from the real world as learning set,

– handling of fuzzy hierarchies,

– efficient handling of aggregate computations, especially for counting.

Hence, the model described in Section 4 for fuzzy multidimensional databases is well adapted. It is used and integrated in the whole system enabling knowledge discovery from fuzzy data. The algorithms used for data mining are described in Section 6. They perform knowledge discovery tasks on fuzzy data stored in fuzzy cubes.

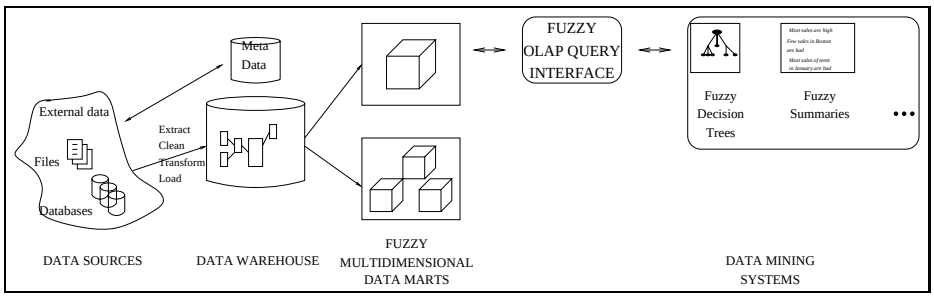

Fig. 5. Architecture

This architecture corresponds to the extension to fuzziness of the architecture proposed for OLAP Mining. Therefore we call it *Fuzzy-OLAP Mining* (or *F-OLAP Mining*). However, we recall that the proposed system can work with crisp data.

6 Data Mining Systems

In the proposed model, several units are available for data mining. Currently, the implemented units concern the generation of fuzzy summaries and the construction of fuzzy decision trees. However, the architecture of the system is modular and is thus open to new units. Current studies concern mainly the construction of fuzzy prototypes, and fuzzy clustering.

Moreover, some units are available for complex queries dedicated to the fuzzy nature of the data that are represented in the database. For instance, one may query the database to discover the distribution of uncertainty degrees regarding a specific dimension. An example of such a query may be *How did the degree of uncertainty on medical diagnosis evolve during this period?*

In this section, we briefly describe each unit that is currently studied. Section 6.1 describes the construction of decision trees while Section 6.2 introduces the process of fuzzy summary generation.

6.1 Fuzzy Decision Trees

Fuzzy decision trees have been presented in Section 3.2. The algorithm to construct fuzzy decision trees is implemented in the Salammbô software ([18]). The output of the system is a fuzzy decision tree that can be considered as a set of discovered classification rules [4].

Constructing fuzzy decision trees from fuzzy multidimensional databases consist in a sequence of OLAP operations of slice and dice and aggregation that we defined above. The algorithm automatically builds fuzzy partitions. This leads to the treatment of numerical data at a higher level of abstraction, and reduces the size of the tree. The whole process of construction is detailed in [17]. The

multidimensional database management system may either send only basic information on data, or on the opposite compute complex fuzzy operations and aggregations, or any of the intermediate solutions.

In the current version of our system, an interface is implemented that consists in exchanging for each node of the tree the statistics on the data associated with the current node. These data may either be singleton values or intervals, which enables us to construct fuzzy decision trees. These statistics are computed using OLAP queries to extract a sub-cube and fetching statistics on it with aggregation functions. These statistics are then used by Salammbô to choose the best attribute to go ahead in the construction of the tree.

In this framework, multidimensional databases are very interesting since they provide efficient tools for aggregation, especially for the computation of the entropy at each step of the construction for the selection of the best attribute to partition data regarding to the class. Moreover, they enable navigation through levels of granularity, thus decision trees can be induced at several levels of granularity.

6.2 Fuzzy Summaries

Fuzzy summaries have been presented in Section 3.2. Many studies have been led on fuzzy summaries, but they mainly concern relational data. Hence, we propose to extend the algorithm for the generation of fuzzy summaries in the framework of the architecture proposed in this paper when considering fuzzy multidimensional databases.

We show here that the needed queries for this process may be processed thanks to fuzzy OLAP queries. In this framework, we take advantage of the efficiency of the multidimensional databases management systems to compute aggregates and to navigate at multiple levels of granularity.

Several types of summaries are studied:

- *simple summaries* - Objects are described by a unique criterion: *"Most sales are bad"*.
- *complex summaries* - Objects are described by several criteria: *"Most sales of shoes in the region of Boston are bad"*.
- *intra-dimension summaries* - Objects are selected regarding a high level of abstraction, and described at a lower level: *"Most good sales results in East are from Boston"*.

The generation of multiple-level summaries has been particularly studied. A user-friendly interface provides the user with tools to generate summaries at the level he wants and to refine the summaries, regarding any dimension he may choose.

Methods to reduce the cost of the generation are studied. However, they are out of the scope of this paper, refer to [16] for a detailed description. These methods are mainly based on the user choices, and on algorithms developed from the generation of association rules ([1]). Some properties that are specific to fuzzy

summaries are also studied. They are based on works on fuzzy quantifiers and fuzzy summarizers.

Furthermore, the generation of fuzzy summaries take advantage of the fuzzy multidimensional database model we define, by taking into account fuzzy hierarchies which enable a value to gradually belong to more than one concept. We suppose that these hierarchies are known before the generation of the summaries, but they also may be automatically built, as it can be done for fuzzy decision trees.

Pretreatments are considered. They provide a way to precompute some results of frequent queries. Moreover, the data marts are built for a specific analysis goal, and the structure of the hypercubes is very important in order to prune summaries that are not of major interest for the user. For instance, the measure of the cube is a dimension of great interest for the analyst and it is the object that has to be explained regarding all the other dimensions of the cube. Thus in the proposed system, the measure is taken as a default object to explain in the summaries. However, the user may change this. The other dimensions are considered depending on whether the user has chosen them or not, and they are taken into account at multiple-level, beginning from the level where the user rolled them up or drilled them down. After each generation of fuzzy summaries, the user is able to consider any dimension at any other level of abstraction. Notice that generating summaries at a high level of abstraction is very efficient and fast, but in many cases, it does not produce relevant information, whereas generating summaries at low levels of granularity is highly time-consuming but it may produce information of greater interest.

7 Implementation

The proposed fuzzy multidimensional database model has been implemented by enhancing the multidimensional database management system `Oracle Express Server` in the *FCUB* system (standing for FUzzy cUBe). This system provides a user-friendly and efficient application to visualize multidimensional data. Imprecise data are shown graphically in order to facilitate data visualization.

The application *FCUB Miner* is also available for data analysis. It provides several data mining units, which are implemented with `Oracle Express Objects`, `Java` and `Microsoft Visual C++`.

Tests have been performed on large amounts of data. Two databases have been used for testing. The first one describes sales regarding to the dimensions $PRODUCT$, $DISTRICT$ and $MONTH$. The second database is a relational database given by the French Department of Education and Research. A multidimensional data mart has been constructed to store results to a French diploma. This multidimensional database contains several cubes. One of them describes the proportion of students that succeeded with honors to this diploma with regards to the total number of students having passed the exam. This information is described regarding to eight dimensions: AGE (44 values), $NATIONALITY$ (2 values), $SOCIAL\_CLASS$ (39 values), $YEAR$ (2 values), SEX (2 values),

$STATUS$ of the high school (7 values, *e.g.* public school, private school), $TYPE$ of the diploma (15 values, *e.g.* science, literature), $OPTION$ (12 values, *e.g.* theater). This cube contains more than 17 million cells, for over 1 million students.

Below are given some of the constructed rules. *young*, *middle age* and *old* are fuzzy labels from a fuzzy partition automatically built by the system.

R1: For *young* students, without major, the proportion of students that pass their exam with honors is *high*.
R2: For *middle age* students, without option, the proportion of students that pass their exam with honors is *bad*.
R3: For *young* students, having option mathematics, for the scientific diploma, the proportion of students that pass their exam with honors is *high*.
R4: For *old* students, for the scientific diploma, the proportion of students that pass their exam with honors is *bad*.

First results concerning the generation of fuzzy summaries can be found in [16].

8 Conclusion

In this paper, a new approach for data mining is described, coupling fuzzy multidimensional databases and fuzzy data mining systems, and achieving knowledge discovery from imperfect data.

An architecture based on fuzzy multidimensional databases is given. It uses these kinds of data repositories to extract relevant knowledge from large data sets from the real world. According to several works that highlighted the great interest of the OLAP framework in the knowledge discovery process, this approach enhances the existing solutions. It provides a way to deal with data from the real world, and to apply flexible operations on data sets stored as multidimensional arrays, generating more understandable fuzzy rules.

Many perspectives are associated with this work. However, our further work concerns mainly:

- the enhancement of the current implementation of the system,
- the integration of other data mining systems (*e.g.,* clustering, prediction),
- the comparison of different approaches managing the complexity of the queries from the data mining system into the multidimensional database management system,
- the use of discovered knowledge as metadata in order to automatically build interesting cubes from the data warehouse,
- the use of user expectations to deal with granularity levels in the discovery process and to enhance the system performances.

References

1. R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In *Proc. of the ACM SIGMOD Conf.*, pp. 207–216, 1993.
2. P. Bosc and J. Kacprzyk, ed.. *Fuzziness in Database Management Systems, Studies in fuzziness.* Springer-Verlag, 1995.
3. P. Bosc, L. Liétard, and O. Pivert. Extended functional dependencies as a basis for linguistic summaries. In *PKDD*, vol. 1510, pp. 93–101. Springer Verlag, 1998.
4. B. Bouchon-Meunier and C. Marsala. Learning fuzzy decision rules. In J. Bezdek, D. Dubois, and H. Prade, ed., *Fuzzy Sets in Approximate Reasoning and Information Systems, The Handbooks on Fuzzy Sets Series*, pp. 279–304. Kluwer Academic, 1999.
5. B. Bouchon-Meunier, C. Marsala, and M. Ramdani. Learning from imperfect data. In D. Dubois, H. Prade, and R. R. Yager, ed., *Fuzzy Information Engineering: a Guided Tour of Applications*, pp 139–148. John Wileys and Sons publisher, 1997.
6. B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards general measures of comparison of objects. *Fuzzy Sets and Systems*, 84:143–153, 1996.
7. E.F. Codd, S.B. Codd, and C.T. Salley. Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate. White Paper, 1993.
8. J.C. Cubero, J.M. Medina, O. Pons, and M.A. Vila. Data summarization with linguistic labels: A loss less decomposition approach. In *IFSA*, 1997.
9. D. Dubois and H. Prade. Fuzzy sets in data summaries - outline of a new approach. In *Proc. of IPMU*, pp. 1035–1040, 2000.
10. J. Han. Olap Mining: An Integration of OLAP with Data Mining. In *IFIP Conf. on Data Semantics*, pp. 1–11, 1997.
11. J. Han. Towards on-line analytical mining in large databases. In *ACM SIGMOD Conf.*, March 1998.
12. J. Han and Y. Fu. Discovery of multiple-level association rules from large databases. *IEEE TKDE*, 11(5):798–804, 1999.
13. J. Kacprzyk, R.R. Yager, and S. Zadrozny. A fuzzy logic based approach to linguistic summaries of databases. *Int. Journal of Applied Mathematics and Computer Science*, 2000.
14. M. Kamber, J. Han, and J.Y. Chiang. Metarule-guided mining of multi-dimensional association rules using data cubes. In *3rd Int. Conf. on Knowledge Discovery and Data Mining (KKD'97)*, 1997.
15. A. Laurent. Bases de données multidimensionnelles floues. *Journées Bases de Données Avancées*, Cepaduès Ed., 2001 (to appear).
16. A. Laurent. Generating fuzzy summaries from fuzzy multidimensional databases. In *Fourth International Symposium on Intelligent Data Analysis, IDA'01.* Springer-Verlag, September 2001, to appear.
17. A. Laurent, B. Bouchon-Meunier, A. Doucet, S. Gancarski, and C. Marsala. Fuzzy data mining from multidimensional databases. *Int. Symp. on Computational Intelligence, Studies in Fuzziness and Soft Computing*, 54:278–283, 2000.
18. C. Marsala and B. Bouchon-Meunier. An adaptable system to construct fuzzy decision trees. In *Proc. of the NAFIPS'99 (North American Fuzzy Information Processing Society)*, pp. 223–227, New York, USA, June 1999.
19. J R. Quinlan. Induction of decision trees. In *Machine learning, volume 1*, pp. 81–106. Kluwer Academic Publishers, 1986.
20. P. Vassiliadis and T. Sellis. A survey of logical models for OLAP databases. *SIGMOD Record*, 28(4), December 1999.