Fuzzy Information Granulation with Multiple Levels of Granularity

Giovanna Castellano, Anna Maria Fanelli, and Corrado Mencar

Abstract. Granular computing is a problem solving paradigm based on information granules, which are conceptual entities derived through a granulation process. Solving a complex problem, via a granular computing approach, means splitting the problem into information granules and handling each granule as a whole. This leads to a multi-level view of information granulation, which permeates human reasoning and has a significant impact in any field involving both human-oriented and machine-oriented problem solving. In this chapter we examine a view of granular computing as a paradigm of human-inspired problem solving and information processing with multiple levels of granularity, with special focus on fuzzy information granulation. To support the importance of granulation with multiple levels, we present a multi-level approach for extracting well-defined and semantically sound fuzzy information granules from numerical data.

Keywords: Fuzzy Information Granulation, Information Granules, Interpretability, Multi-level Granulation, Conditional Fuzzy C-Means, Double Clustering.

1 Introduction

In the last decade, information granulation has emerged as a powerful tool for data analysis and information processing, which is in line with the way humans adopt to process information. We perceive the world by structuring our knowledge, perceptions, and acquired evidence in terms of information granules that offer abstractions of the complex world and phenomena. Being abstract constructs, information granules and the processing of them, referred to as Granular Computing (GrC), provide problem solvers with a conceptual and algorithmic framework to deal with several real-world problems.

The term GrC spans a variety of disciplines, thus it is often loosely defined as an umbrella term covering any theories, methodologies, techniques, and tools that

Giovanna Castellano · Anna Maria Fanelli · Corrado Mencar
Department of Informatics, University of Bari “Aldo Moro”
Via Orabona, 4 – 70125 Bari, Italy

W. Pedrycz and S.-M. Chen (Eds.): Granular Computing and Intell. Sys., ISRL 13, pp. 185–202.
springerlink.com

make use of granules in complex problem solving (Yao 2005). Various frameworks for information granulation have been proposed so far, with a growing diversity of formalisms. One of the main formalism proposed for GrC is the theory of fuzzy sets (Zadeh 1997) introduced by Zadeh, who considers granular computing as basis for computing with words, i.e., computation with information described in natural language (Zadeh 1997; Zadeh 2006). Besides, many other formalisms can be considered fundamental for GrC, stemming from rough sets (Pawlak 1998; Yao 2001; Peters et al. 2003), interval analysis (Bargiela and Pedrycz 2003), shadowed sets (Pedrycz 2005), etc.

Beyond the theory underlying a GrC framework, two main features are desirable for any information granulation approach:

1. the ability to determine the granularity level that better represents the nature of data;
2. the ability to provide users with information granules that both represent the data accurately and carry a clear semantic meaning, i.e. granules that are *interpretable* for human users.

These features bring a shift in GrC, from a paradigm of machine-centric information processing to a paradigm of human-inspired problem solving. This shift is considered one of the recent trends in GrC research (Yao 2008). In this endeavor, it is well accepted that the theory of fuzzy sets has largely contributed to the emergence of granular computing as a paradigm of human-inspired problem solving and information processing (Bortolan and Perdrycz 2002). Actually, human-centered information processing was initiated with the introduction of fuzzy sets, which successively led to the development of the GrC paradigm (Bargiela and Pedrycz 2008; Zadeh 1997). The object of fuzzy information granulation is to build models by means of information granules that are quantified in terms of fuzzy sets, i.e. conceptual entities with well-defined semantics that are interpretable by humans.

To stress the shift from machine-centric computing toward human-inspired computing, a common theme is converging to a view of GrC as a paradigm of information processing with an underlying notion of multiple levels of granularity. The introduction of multiple levels of granularity corresponds to consider multiple levels of abstractions for a given problem, with each level capturing particular aspects of the problem. To some extent, this may avoid limitations of a single level of representation.

Along with these ideas, in this chapter we examine the current research directions on multi-level granular computing, with special focus on fuzzy information granulation. To give evidence of the advantage of multi-level granulation, we present an approach to perform fuzzy granulation with multiple levels of granularity in a hierarchical fashion. In particular, a framework of representation with two levels of granularity is described. At the first level, the whole dataset is granulated. At the second level, data embraced in each first-level granule are further granulated taking into account the context generated by that granule. The derived hierarchical collection of granules can be used to construct a committee of fuzzy models providing a good balance between interpretable representation and precise approximation.

2 Multi-level Granular Computing

The basic components of GrC are information granules, intended as chunks of knowledge made of different objects “drawn together by indistinguishability, similarity, proximity or functionality” (Zadeh 1997). Information granules yield an abstraction of the reality in form of concepts depending on the context. For such reason, a granulation process, i.e. the process of constructing granules, serves not only to design a model, but also to simplify our understanding of it. As computing units, granules can decompose a complex problem into some simpler or smaller problems, so that the computing costs are reduced and the problem can be better understood.

Based on complexity, abstraction grade and size, granules can be evaluated at different levels. The problem domain, i.e., the universe of discourse, represents the coarsest granule at the highest level. Granules at the lowest level are composed of elements or basic particles of the particular model that is used (Yao, 2007). According to the level of granularity taken into account, i.e. to the point of view, a granule “may be an element of another granule and is considered to be a part forming the other granule. It may also consist of a family of granules and is considered to be a whole” (Yao 2008). This leads to the notion of levels. While each granule provides a local view, a level provides a global view.

An important property of granules and levels is their granularity, i.e. the size of information granules and their distribution. By changing the granularity, we can control the amount of details so as to hide or reveal more or less details about the problem at hand (Bargiela and Pedrycz 2003). The lower level of granularity can yield the most detailed information, but some useful knowledge may be buried into unnecessary details. On the other hand, the higher level of granularity might reduce some information, but it can provide users with a better insight into the essence of data. From one level we can pass to a lower level of granularity by means of a granulation process that decomposes a whole into parts; this corresponds to “analytical thinking” whereas, going to an upper level merging parts into wholes, corresponds to “synthetic thinking” (Yao 2007).

Granularity enables us to properly arrange granules and levels, so as to derive a hierarchical view of the problem at hand. In building a hierarchical structure, we discover a vertical separation of levels and a horizontal separation of granules at the same hierarchical level (Yao 2009). Usually, the two separations must ignore information that is irrelevant to the current interest or does not greatly affect our solution. Furthermore, a single hierarchy only represents one view. As illustrated by Yao, granular structures enable both a multi-level view (given by a single hierarchy) and a multi-view understanding (given by many hierarchies) (Yao 2009). The latter stresses the consideration of diversity in modelling, for which we look at the same problem from many perspectives. This is useful when, in order to understand a problem, we need to explore multiple representations of it, in terms of multiple views and multiple levels of abstraction.

Summarizing, a fundamental key to GrC is representing and working with different levels of granularity in every stage of problem solving. We may view a problem through many different facets, and associate a representation with a particular view capturing specific aspects of the problem. For each view, we may consider multiple levels of abstractions, each representing the problem at a particular level of details. This particular issue of granulating information by using high-order properties, i.e. properties of collections of granules formed at higher levels, lies behind every - not necessarily computational - task involving problem solving: it describes a way of thinking, named “granular reasoning” (Zhong at al. 2008), that relies on the human ability to perceive the real world under various levels of granularity. In (Yao 2006), some fundamental issues related to the notion of multiple levels of granularity are addressed and analyzed from three different perspectives, namely philosophical, methodological and computational perspective leading to the so-called *triarchic theory* of GrC (Yao 2005; Yao 2007).

The philosophical facet of GrC offers a worldview in terms of structures as represented by multiple levels. This leads to a way of structured thinking, made of levels of abstraction, which is applicable to many branches of natural and social sciences. We may consider, for example, levels of understanding in education, levels of interpretation in history and language understanding, levels of organization in ecology and social sciences, levels of processing in modeling human memory, and many others. For example, Jeffries and Ransford proposed a multiple hierarchy model to integrate class, ethnicity, gender, and age for the study of social stratification (Jeffries and Ransford 1980). They show how the traditional single hierarchy approach based on social classes limits one’s understanding of the complexities of modern societies, while a multiple hierarchy approach could increase one’s understanding and be more comprehensive and valid for studying social inequality.

The methodological perspective of GrC raises quite natural. As a general method of structured problem solving, GrC provides practical strategies and effective principles that are used by humans for solving real-world problems. Those principles of granular computing have, in fact, been extensively used in different disciplines under different names and notations. For example, many principles of structured programming or software design can be readily adopted for granular computing (Han and Dong 2007). As another example, in (Belkhouche and Lemus-Olalde 2000) an abstract interpretation of multiple views in software design is formalized. In the process of modeling a system, the designer always generates a set of designs, such as functional, behavioral, structural and data designs. Each design focuses on a view that describes a subset of relevant features of a system and is expressed by one or more notations. The authors argued that a multiple view analysis framework can be used to systematically compare, identify and analyze the discrepancies among different views, enhance design quality and provide a multi-angled understanding of a problem or a project.

The computational perspective of GrC underlies the other two if we consider the term “computing” in its broad meaning to include information processing in the abstract, in the brain and in machines. As a paradigm of structured information processing, granular computing focuses on computing methods based on granular structures. In particular, when computation is intended as information processing by machines, i.e. data analysis, the use of hierarchical granular structures leads to multi-view intelligent data analysis, which explores data from different perspectives to reveal various types of structures and knowledge embedded in the data. Each view may capture a specific aspect of the data and hence satisfy the needs of a particular group of users. Collectively, multiple views provide a comprehensive description and understanding of the data. According to this idea, Chen and Yao proposed a multi-view approach that provides a unified framework for integrating multiple views of intelligent data analysis (Chen and Yao 2008). Managing multiple representations of the same data at different levels of granularity is widely recognized as a relevant one also in the community of spatial database. In (De Fent et al. 2005) spatial data are modeled through conceptual models with two distinct, but related, granularity dimensions: a spatial one and a semantic one. In this case, a multi-level granulation enables to capture scale as well as semantic changes and to constrain the relationships between them.

From a different viewpoint, we may say that a multi-level information granulation approach turns out to be beneficial as soon as we need to represent and solve the problem through a set of linguistic concepts. Actually, multi-level granulation has been widely investigated in the realm of linguistic approaches, i.e. approaches using linguistic terms to represent information in a qualitative way (Glöckner and Knöll 2001). Linguistic representations, such as those based on fuzzy sets, are especially suitable when information is unquantifiable due to its imprecise nature (e.g., when evaluating the “comfort" of a car, only terms like “good", “fair", “poor" can be used), but they can be used as well when a quantitative representation of information cannot be stated because either it is unavailable or the cost for its computation is too high, thus an approximate value can be tolerated (e.g., when evaluating the speed of a car, linguistic terms like “fast", “very fast", “slow" can be used instead of numeric values). When using a linguistic representation of information, an important parameter to determine is the semantic granularity, i.e. the cardinality of the linguistic term set used to express the information. According to the uncertainty degree of a domain expert, the linguistic term set used to represent her knowledge may have more or less terms. When different experts have different uncertainty degrees in their knowledge, then several linguistic term sets with different granularities are necessary. To cope with this multiple source of uncertain information, multi-granularity linguistic term sets based on fuzzy set theory have been proposed and applied in several fields, such as decision making (Herrera et al. 2000; Mata et al. 2009) and information retrieval (Herrera-Viedma et al. 2003).

Multi-level GrC has been also applied to represent taxonomies of concepts. In (Qiu et al. 2007) a hierarchy of granules corresponding to ontological concepts is built by an information table using rough-set techniques. A granular space model for ontology learning is explored, to describe domain ontologies at different granularities and hierarchies. In (Gu et al. 2006) an approach for constructing hierarchy of granules based on fuzzy concept lattices is proposed. The knowledge granularity is discussed, and an algorithm for constructing a hierarchical structure of coarser granules is also illustrated.

The potential applications and implications of multi-level granulation can be also recognized in the area of Web intelligence. As shown in (Yao 2007), GrC may provide the necessary theory for designing and implementing new types of web-based information processing systems based on a conceptual model of the human brain. Yao and Yao discuss how Web information retrieval can benefit from grouped and personalized views provided by a GrC process (Yao and Yao 2003). The use of both single-level and multi-level granulations of web documents is investigated. In (Li et al. 2001) a technique for automatically constructing multi-granular and topic-focused site maps using trained rules on Web page URLs, contents, and link structure is presented. This type of multi-granular site maps can support better interaction for users to scale up and scale down the details. The system provides more detail on the regions relevant to the focused topic while keeping the rest of the map compact so that the users can visualize their current navigation positions relative to other landmark nodes in the Web site.

To conclude, we emphasize the crucial role of interpretability in the realm of GrC. As a paradigm for human-centric information processing, GrC should provide a common interface for communicating information between humans and machines. This human-oriented communication is partially achieved by representing perceptual information in a computer manageable form, i.e. by means of information granules. To make more effective this communication, information granules should be interpretable, i.e. semantically co-intensive with human knowledge. Interpretability of information granules is a complex requirement that needs a comprehensive analysis of all facets of the problem for which granules are developed and used. Therefore interpretable information granulation opens several methodological issues, regarding the representation and manipulation of information granules, the interpretability constraints and the granulation processes (Mencar and Fanelli 2008; Mencar 2009). We believe that addressing all such issues at multiple levels of granularity leads to a sight of GrC as an effective tool to design information processing systems characterized by a strong human-centric imprint. To support this idea, in the next section we present a multi-level GrC strategy to derive interpretable information granules from data.

3 A Multi-level Approach for Information Granulation

As an example of multi-level granulation strategy, in this section we describe a multi-level approach for fuzzy information granulation. The approach is based on DC*f* (Double Clustering framework), a framework to create interpretable granules from data by taking into account a number of interpretability constraints (Castellano et al. 2005). DC*f* extracts fuzzy information granules that can be easily labeled with semantically sound linguistic labels. Moreover, the number of information granules to be extracted can be kept small, so as to provide a compact (and hence readable) description of data. Nevertheless, DC*f* provides for a flat representation of information granules, and the user is committed to define the granularity level. This is due to the fact that DC*f* can extract a fixed number of information granules. A high number of information granules leads to an accurate yet unreadable description of data, while a small number of information granules provides a highly interpretable description of data, but the employment of such granules in fuzzy predictive models may not result satisfactory because of a possibly coarse accuracy.

In order to obtain a more accurate description of data while keeping interpretability, we have developed an extension of DC*f*, called ML-DC (Multi-Level Double Clustering), which is intended to provide a multi-level granulation of available data, i.e. a granulation of data at different levels, in a hierarchical fashion (Castellano et al. 2007). ML-DC exploits the DC*f* structure to provide for multiple view of data: a coarse qualitative view where main relationships are described by large granules, and a more refined, quantitative granulation that could be used for defining the predictive model. At the first level, the whole dataset is granulated, while, at the second level, data embraced in each first-level granule are further granulated taking into account the context generated by that granule. Based on the extracted multi-level granules, a hierarchical committee of Fuzzy Inference Systems (FISs) is constructed that can approximate a mapping with a good balance between accuracy and interpretability.

Roughly speaking, ML-DC operates information granulation at two levels:

- at first level, granulation of data is carried out according to a specific granularity, as in DC*f;*
- at second level, for each discovered information granule, data are re-aggregated for a further granulation process.

The process can be reiterated for a number of levels. However a two-level granulation is adequate to obtain two views of the problem (a qualitative one and a quantitative one), so as to achieve a balanced trade-off between accuracy and interpretability of data. Hence, through the application of ML-DC, two levels of fuzzy information granules are built from data: granules of the first level are used to roughly describe data through qualitative linguistic labels; granules of the second level are used to describe each information granule of the first level. This is aimed at finding a more accurate description of the hidden relationships lying among data and – a the same time – preserve interpretability of the extracted knowledge since interpretability constraints are satisfied for both levels of granulation.

The first-level granulation in ML-DC is made according to the double-clustering framework DC*f*, which involves two main steps:

1. multi-dimensional clustering on the whole dataset, providing a collection of multidimensional prototypes;
2. one-dimensional clustering of the projection of the derived prototypes along each dimension, yielding to one-dimensional prototypes.

In ML-DC we perform step 1 by means of fuzzy clustering and step 2 by means of hierarchical clustering. Fuzzification of the information granules is achieved by first fuzzifying the one-dimensional granules defined by the one-dimensional prototypes and then by aggregating one-dimensional fuzzy sets to form multi-dimensional fuzzy information granules. For each dimension, the extracted clusters are transformed into as many interpretable fuzzy sets. Fuzzy sets with

Gaussian membership function are considered, whose centers and widths are defined so as to take into account the information provided by the clustering stages and, at the same time, to meet the following interpretability constraints:

- *Normality*, *convexity* and *continuity*: these constraints are verified as soon as Gaussian membership functions are chosen for fuzzy sets;
- *Proper ordering*: this is verified by defining the order relation of fuzzy sets reflecting the order of the prototypes.
- *Justifiable number of elements*: this constraint is verified by an appropriate choice of the number of prototypes;
- *Distinguishability* and *completeness*: these constraints are verified by the construction of the fuzzy sets, which is made so as to not exceed an overlap threshold ε and to guarantee the ε-coverage.

The second-level granulation is carried out according to the same double-clustering schema, but taking into account the context generated from each first-level information granule. Indeed, if this context is ignored, the second-level granulation would be identical to first-level granulation and no additional information would be derived from data. To keep into account contextual information in the second-level granulation, multi-dimensional clustering is performed by the Conditional Fuzzy C-Means (CFCM) proposed by Pedrycz, which is an extension of the well-known FCM clustering algorithm (Pedrycz 1996). The CFCM clustering algorithm minimizes the following objective function:

$$J\left(\mathbf{U},\mathbf{V}\right)=\sum_{j=1}^{c}\sum_{i=1}^{N}u_{ij}^{m}\left\|\mathbf{x}_{i}-\mathbf{v}_{j}\right\|^{2} \tag{1}$$

where $\mathbf{x}_i \in \mathbb{R}^n$, $i=1,2,\ldots,N$ are the observational n-dimensional data to be clustered, c is the number of clusters, $\mathbf{U}=\left[u_{ij}\right]_{j=1,2,\ldots c}^{i=1,2,\ldots,N} \in [0,1]^{N\times c}$ is the partition matrix, being u_{ij} the membership degree of the i-th observation to the j-th cluster, $\mathbf{V}=\left[\mathbf{v}_j\right]_{j=1,2,\ldots,c} \in \mathbb{R}^{n\times c}$, is the matrix of the prototypes corresponding to the fuzzy clusters and m is a fuzzification parameter, here fixed to 2.0.

The objective function (1) is constrained as follows to avoid degenerate solutions:

$$\forall j=1,2,\ldots,c:\sum_{i=1}^{N}u_{ij}>0 \tag{2}$$

and:

$$\forall i=1,2,\ldots,N:\sum_{j=1}^{c}u_{ij}=f_i \tag{3}$$

Constraint (3) differentiates CFCM from the standard FCM clustering algorithm. Such constraint requires that the sum of memberships of a point to each cluster is equal to a constant $f_i \in [0,1]$ that defines the *context* of the clustering process. For the first-level granulation process, no context is defined for ML-DC. Hence, CFCM is reduced to standard FCM by setting:

$$\forall i = 1,2,\ldots,N : f_i = 1 \tag{4}$$

For the second-level granulation process, a context is defined by each first-level fuzzy information granule:

$$\forall i = 1,2,\ldots,N : f_i = \mu_k^{\mathrm{I}}(\mathbf{x}_i) \tag{5}$$

where $\mu_k^{\mathrm{I}}(\mathbf{x}_i)$ is the membership degree of the *i*-th observation to the *k*-th fuzzy information granule discovered in the first-level granulation process. The quantities f_i establish the connection between first-level and second-level information granules. Information granules at the second level are indeed forced to focus their location in the fuzzy sub-region of the domain where each first-level information granule is placed.

The CFCM clustering algorithm follows the Alternating Optimization strategy for the minimization of the objective function (1). This strategy is iterative. At each iteration the prototypes and the partition matrix are updated according to the following formulas:

$$\mathbf{v}_j[\tau+1] = \frac{\sum_{i=1}^{N} u_{ij}^m[\tau]\mathbf{x}_i}{\sum_{i=1}^{N} u_{ij}^m[\tau]} \tag{6}$$

and:

$$u_{ij}[\tau+1] = \frac{f_i}{\left(\sum_{k=1}^{c} \frac{\left\|\mathbf{x}_i - \mathbf{v}_i\right\|^2}{\left\|\mathbf{x}_i - \mathbf{v}_j\right\|^2}\right)^{\frac{1}{m-1}}} \tag{7}$$

Summarizing, ML-DC is a double-clustering approach where CFCM is used in the multi-dimensional clustering and hierarchical clustering is used in the one-dimensional clustering. For first-level granulation, ML-DC is applied to data with constant context. The result of first-level granulation is a set of K^{I} fuzzy information granules G_k^{I} with membership functions $\mu_k^{\mathrm{I}} : \mathbb{R}^n \to [0,1]$, $k = 1,2,\ldots,K^{\mathrm{I}}$.

ML-DC exploits the advantages of multi-level information granulation to reach a tradeoff between accuracy and interpretability. Here we describe a granulation methodology based on ML-DC, which is specifically designed for classification problems, even though the extension to function approximation problems is straightforward. Let $\mathbf{X} \subseteq \mathbb{R}^n$ be the Universe of Discourse of data (assumed as hyper-box) and $\mathbf{C} = \{1,2,\ldots,C\}$ a set of class labels. Let

$D=\{\langle \mathbf{x}_i, c_i\rangle \in \mathbf{X}\times\mathbf{C}, i=1,2,\ldots,N\}$ a set of pre-classified observational data. The first-level granulation of data provides for K^I fuzzy information granules with membership functions μ_k^{I}, $k=1,2,\ldots,K^I$. The value K^I can be fixed as small as desired (e.g. less than 7±2 according to (Valente de Oliveira 1999)). As shown above, each information granule G_k^I satisfies a number of interpretability constraints so that it can be labeled linguistically.

For each G_k^I and each class label $c\in\mathbf{C}$, we compute the relative frequency of observations of class c belonging to the fuzzy granule G_k^I, as follows:

$$\pi_{k,c}^I = \frac{\sum_{i=1,c_i=c}^{N} \mu_k^I(\mathbf{x}_i)}{\sum_{i=1}^{N} \mu_k^I(\mathbf{x}_i)} \tag{8}$$

These values can be used to build a set of K^I fuzzy rules with the following schema:

$$\begin{aligned} \text{IF}\,\mathbf{x}\,\text{is}\,G_k^I\ \text{THEN}\ & P(class=1)=\pi_{k,1}^I \\ & P(class=2)=\pi_{k,2}^I \\ & \cdots \\ & P(class=C)=\pi_{k,C}^I \end{aligned} \tag{9}$$

Given an input $\mathbf{x}$, the outputs of the classifier are computed according to the following formula:

$$\pi_c^I(\mathbf{x}) = \frac{\sum_{k=1}^{K^I} \mu_k^I(\mathbf{x})\pi_{k,c}^I}{\sum_{k=1}^{K^I} \mu_k^I(\mathbf{x})} \tag{10}$$

for $c=1,2,\ldots,C$. If one class only must be assigned to $\mathbf{x}$, then the class with highest $\pi_c^I(\mathbf{x})$ is chosen (tiers are selected arbitrarily).

The FIS classifier designed through first-level information granulation is very compact (and hence highly interpretable) but expectedly not very accurate. Second-level information granules can be exploited to improve accuracy of the estimated mapping as follows. For each first-level information granule, ML-DC provides a set of second-level information granules that can be used to generate a corresponding FIS with the same rule schema of the first-level FIS. As a result, a set of K^{I} FISs are generated. All such FISs are connected to form a hierarchical committee of FISs from which the input/output mapping is inferred. The outputs of the hierarchical FIS are defined as the weighted sum of the outputs of each FIS belonging to the committee. Formally, given an input $\mathbf{x}\in\mathbb{R}^n$, the output of the FIS committee is:

$$\pi_c^{II}(\mathbf{x}) = \frac{\sum_{k=1}^{K^{\mathrm{I}}} \pi_{k,c}^{II}(\mathbf{x})\, w_k}{\sum_{k=1}^{K^{\mathrm{I}}} w_k} \tag{11}$$

where $\pi_{k,c}^{II}(\mathbf{x})$ is the output of the k-th FIS belonging to the committee relatively to class c, while w_k is the weight assigned to the k-th FIS, corresponding to the degree of membership of the input $\mathbf{x}$ to the antecedent part of the first-level information granule:

$$w_k = \mu_{X,k}^{\mathrm{I}}(\mathbf{x}) \tag{12}$$

In this way, the weight of a FIS in the committee is high (and hence the corresponding output is very relevant in determining the final output) when the input falls within the associated first-level information granule. On the contrary, the weight becomes as small as far the input is from the first-level granule prototype.

In summary, two models are derived from the application of ML-DC (Fig. 2). A simple FIS generated from first-level granulation that can be used mostly for representation purposes, and a hierarchical committee of different FISs that can be used to model the input/output mapping. In this way, the trade-off between accuracy and interpretability can be well balanced.

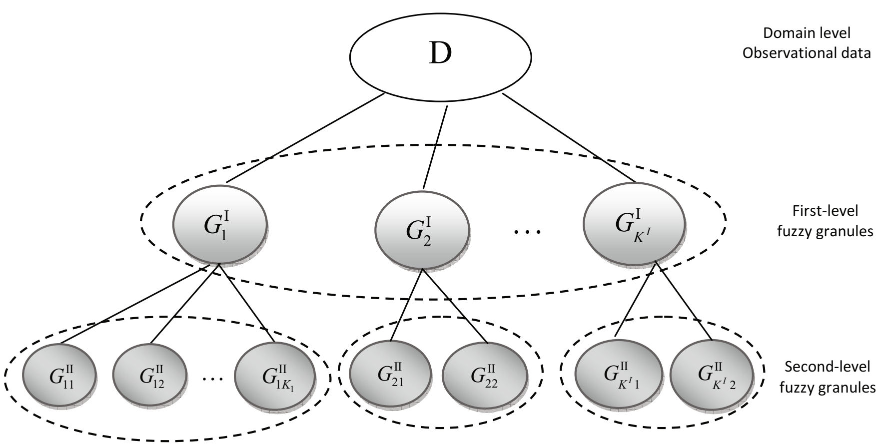

Fig. 1 The multi-level granulation obtained by ML-DC.

A Illustrative Example

In this section we report some simulation results from the application of ML-DC to solve the benchmark classification problem provided by the Cone-Torus (CT) dataset[1]. The CT dataset is available as a collection of a training set and a test set,

[1] http://www.bangor.ac.uk/~kuncheva/Z.txt

both consisting of 400 examples. Each example is two-dimensional and labeled with one out of three class labels. Data are generated synthetically so as to provide the distribution depicted in Fig. 2. This classification problem is inherently difficult, because of the nonlinear class boundaries and high overlapping of data belonging to different classes. According to the literature, simple classifiers are unable to provide an acceptable classification accuracy (classification error about 20-30%), while more complex models (such as FISs with more than 100 rules) provide a classification error not below 10% about (Kuncheva 2000). As a consequence, it is very hard to provide both a linguistically interpretable and accurate description of the Cone-Torus dataset.

ML-DC was applied to the data to perform a two-level granulation. At both levels of granulation, the number of information granules to be generated is not greater than five. Starting from the first-level granules, a FIS classifier was derived, while a committee of FIS classifiers was derived from the second-level granules. The rule base obtained for the first-level FIS is reported in Table 1, while the fuzzy sets generated by the granulation procedure are depicted in Fig. 3. At this level, information granules are quite rough, hence qualitative linguistic terms are more appropriate to represent knowledge. Fuzzy inference with this FIS provides a classification error of 28.25% on the training set and 26.75% on the test set. While the fuzzy model is highly interpretable, its classification ability is quite rough, especially in comparison with other black-box models known in literature.

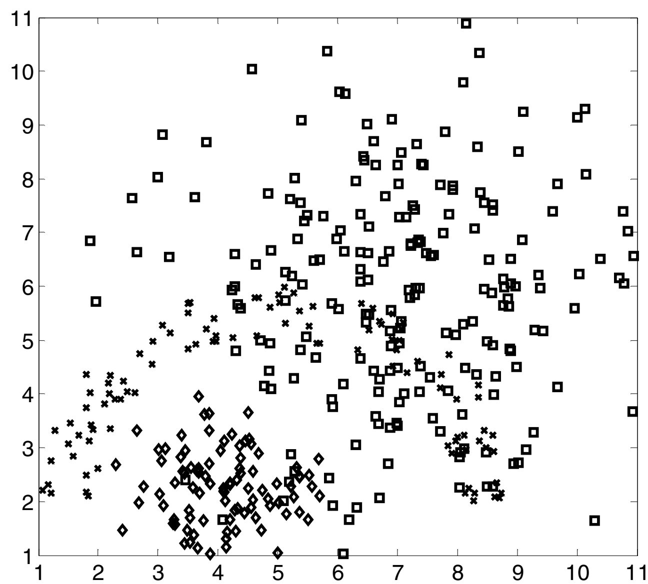

Fig. 2 The Cone-Torus dataset. Class labels are represented as diamonds, squares and crosses.

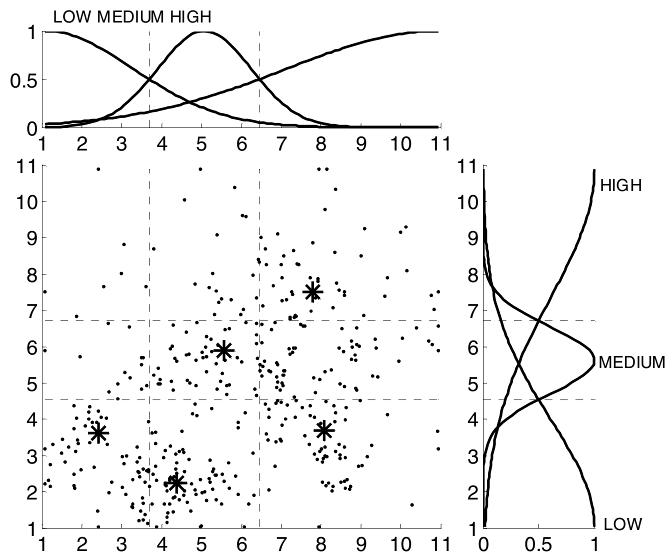

Fig. 3 First-level fuzzy information granulation. Stars represent multi-dimensional prototypes

Table 1 The fuzzy rule base of the first-level FIS

```
IF x is LOW AND y is LOW THEN P(class=1)=0.53, P(class=2)=0.38, P(class=3)=0.09
IF x is MEDIUM AND y is LOW THEN P(class=1)=0.6, P(class=2)=0.11, P(class=3)=0.29
IF x is MEDIUM AND y is MEDIUM THEN P(class=1)=0.01, P(class=2)=0.45, P(class=3)=0.54
IF x is HIGH AND y is LOW THEN P(class=1)=0.21, P(class=2)=0.27, P(class=3)=0.52
IF x is HIGH AND y is HIGH THEN P(class=1)=0.02, P(class=2)=0.11, P(class=3)=0.87
```

The committee of FISs derived from the second-level granulation process provides a classification error of 17.25% on the training set and 13.34% on the test set. These results are comparable with other models known in literature, as reported in Table 2. Each FIS in the committee describes a context, i.e. an information granule derived in the first-level granulation process. Because of their finer granularity, second-level information granules are labeled as fuzzy quantities. In Table 3, the rule-base of the second-level FIS derived for the context “x is LOW and y is LOW” is reported, while the fuzzy sets derived by the granulation procedure are depicted in Fig. 4. It should be noted that for each dimension, leftmost and rightmost fuzzy sets have constant membership value one for elements outside the context. This is in coherence with the semantics of the linguistic terms associated to these fuzzy sets.

Table 2 Classification results for the Cone-Torus dataset of different models known in literature

Model	Class. error on the training set	Class. error on the test set
Nearest mean	29.50%	26.25%
Linear Discriminant	25.50%	23.00%
Quadratic Discriminant	19.50%	16.75%
Parzen	8.75%	12.25%
Nearest Neighbor	17.25%	15.25%
Multi-Layer Perceptron (2-15-3)	13.50%	12.00%
LVQ1 (20 prototypes)	15.50%	14.50%
Wang-Mendel (10 fuzzy sets per input 100 rules)	13.50%	13.00%
ML-DC	17,25%	13,34%

Table 3 The rule base of a second-level FIS

```
Context: x is LOW and y is LOW
IF x is about 1.1 or less AND y is about 2.6 THEN P(class=1)=0.31, P(class=2)=0.38, P(class=3)=0.01
IF x is about 1.1 or less AND y is about 4.5 or more THEN P(class=1)=0.12, P(class=2)=0.88,
P(class=3)=0.00
IF x is about 3.7 or more AND y is about 1.0 or less THEN P(class=1)=0.96, P(class=2)=0.02,
P(class=3)=0.02
IF x is about 3.7 or more AND y is about 2.6 THEN P(class=1)=0.88, P(class=2)=0.06, P(class=3)=0.05
```

Based on the respective features of the two systems, the first-level FIS and the committee of second-level FISs can be used synergistically for tackling complex problems: the first-level FIS can be used to provide a rough description of data, with the primary focus of providing a first understanding of the hidden relationships laying among data. The committee of second-level FISs can be effectively used to accurately classify patterns. Each FIS in the committee locally describes a piece of the Universe of Discourse in an interpretable fashion, so as to provide a more detailed representation of the acquired knowledge. The locality of the description is determined by the contexts obtained by the first-level granulation process. As a result, the interpretability of the overall system is preserved.

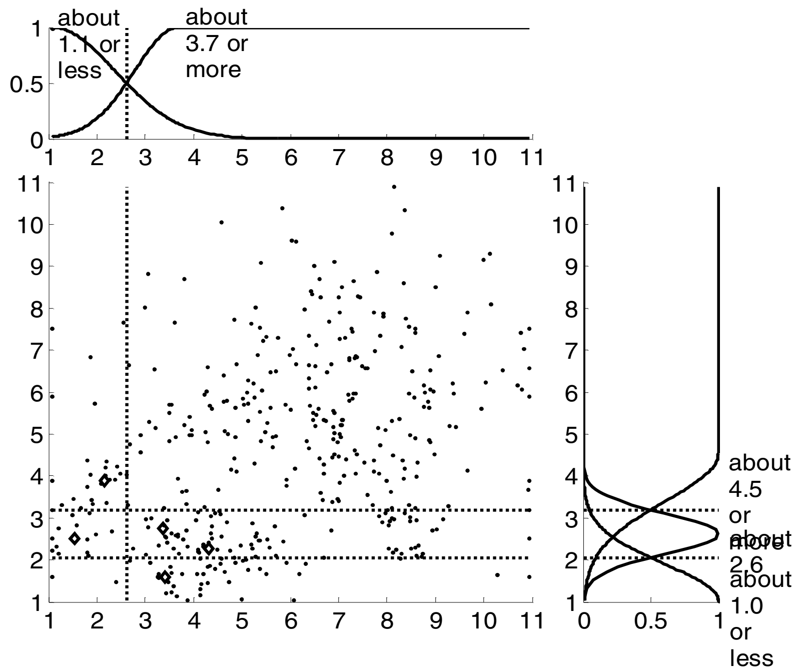

Fig. 4 Second-level information granules for the context "x is low and y is low". Diamonds represent multidimensional prototypes

4 Conclusions

The necessity and the benefits of using GrC in solving real-world problems have been widely pointed out in the recent literature, with an increasing proposal of concrete models and innovative methodologies. As a philosophy and a general methodology, GrC empowers problem solving in many fields; as a paradigm of structured information processing, it supports the development of human-inspired information processing systems. The chance of dealing with different levels of granularity in every stage of problem solving, makes GrC a powerful paradigm for representing and solving problems by means of multi-level strategies. A multi-level granulation process provides several granulated views of the same problem, enabling the focus on useful information structures without looking into too much details.

In this chapter we have emphasized the new perspective of multi-level GrC, as a way to better represent and understand knowledge. As an example of multi-level granulation strategy that improves understanding of data, we have presented ML-DC, a framework to perform a multi-level granulation of data with a balance between accuracy and interpretability. A complex problem is sliced into contexts, which can be used to provide a high-level –yet highly interpretable– description of available data. The contexts are used to locally acquire more accurate fuzzy models, which still preserve interpretability constraints for a readable representation of knowledge. The resulting structure of granules can provide a comprehensible as well as accurate knowledge base. In principle, MD-DC can derive more than two levels of granulation. However, the deeper the granulation level is, the more difficult the comprehensibility of the acquired knowledge. Two levels are deemed enough for a good balance between interpretability and accuracy, since the first level describes the data from a qualitative view, while the second level provides a more quantitative description of the data. Anyway three or more levels might be considered according to the specific application domain.

References

Bargiela, A., Pedrycz, W.: Recursive information granulation: aggregation and interpretation issues. IEEE T. Syts. Man Cyb. 33, 96–112 (2003)

Bargiela, A., Pedrycz, W.: Toward a theory of granular computing for human-centered information processing. IEEE T. Fuzzy Sys. 16, 320–330 (2008)

Belkhouche, B., Lemus-Olalde, C.: Multiple views analysis of software designs. Int. J. Softw. Eng. Know. 10, 557–579 (2000)

Bortolan, G., Pedrycz, W.: Fuzzy descriptive models: an interactive framework of information granulation. IEEE T. Fuzzy Sys. 10(6), 743–755 (2002)

Castellano, G., Fanelli, A.M., Mencar, C.: DCf: A Double Clustering framework for fuzzy information granulation. In: Proceedings of the IEEE International Conference on Granular Computing, Beijing, China, pp. 397–400 (2005)

Chen, Y.H., Yao, Y.Y.: A multiview approach for intelligent data analysis based on data operators. Inform. Sciences 178(1), 1–20 (2008)

De Fent, I., Gubiani, D., Montanari, A.: Granular GeoGraph: a Multi-Granular Conceptual Model for Spatial data. In: Proceedings of 13th Italian Symposium on Advanced Database Systems, Bressanone, Italy, pp. 368–379 (2005)

Glöckner, I., Knöll, A.: A formal theory of fuzzy natural language quantification and its role in granular computing. In: Pedrycz, W. (ed.) Granular Computing: An Emerging Paradigm. Physica-Verlag, Wurzburg (2001)

Gu, S.M., Zhu, S.X., Ye, Q.H.: An Approach for Constructing Hierarchy of Granules Based on Fuzzy Concept Lattices. In: Proceedings of the 5th International Conference on Fuzzy Systems and Knowledge Discovery, Jinan, China, pp. 679–684 (2006)

Han, J., Dong, J.: Perspectives of granular computing in Software Engineering. In: Proceedings of the IEEE International Conference on Granular Computing, pp. 66–71. Silicon Valley, California (2007)

Herrera, F., Herrera-Viedma, E., Martínez, L.: A fusion approach for managing multi-granularity linguistic term sets in decision making. Fuzzy Set Syst. 114, 43–58 (2000)

Herrera-Viedma, E., Cordòn, O., Luque, M., Lopez, A.G., Muñoz, A.M.: A model of fuzzy linguistic IRS based on multi-granular linguistic information. Int. J. Approx. Reason. 34(2-3), 221–239 (2003)

Jeffries, V., Ransford, H.E.: Social Stratification: A Multiple Hierarchy Approach. Allyn and Bacon Inc., Boston (1980)

Kuncheva, L.I.: Fuzzy Classifier Design. Physica Verlag, Berlin (2000)

Liang, J.Y., Li, D.Y.: Uncertainty and Knowledge Acquisition in Information Systems. Science Press, Beijing (2005)

Li, W.S., Ayan, N.F., Kolak, O., Vu, Q., Takano, H.: Constructing multi-granular and topic-focused web site maps. In: Proceedings of the 10th International Conference on the World Wide Web, Hong Kong, pp. 343–354 (2001)

Liang, J.Y., Shi, Z.Z.: The information entropy, rough entropy and knowledge granulation in rough set theory. Int. J. Uncertain Fuzz 12(1), 37–46 (2004)

Mata, F., Martınez, L., Herrera-Viedma, E.: An Adaptive Consensus Support Model for Group Decision-Making Problems in a Multigranular Fuzzy Linguistic Context. IEEE T. Fuzzy Syst. 17(2), 279–290 (2009)

Mencar, C.: Interpretability of Fuzzy Information Granules. In: Bargiela, A., Pedrycz, W. (eds.) Human-Centric Information Processing Through Granular Modelling. Springer, Berlin (2010)

Mencar, C., Castellano, G., Fanelli, A.M.: Balancing interpretability and accuracy by multi-level fuzzy information granulation. In: Proceedings of the IEEE International Conference on Fuzzy Systems, Vancouver, Canada, pp. 2157–2163 (2006)

Mencar, C., Fanelli, A.M.: Interpretability constraints for fuzzy information granulation. Inform. Sciences 178(24), 4585–4618 (2008)

Pawlak, Z.: Granularity of knowledge, indiscernibility and rough sets. In: Proceedings of the IEEE International Conference on Fuzzy Systems, pp. 106–110. Piscataway, New Jersey (1998)

Pedrycz, W.: Conditional Fuzzy C-Means. Pattern Recogn. Lett. 17, 625–632 (1996)

Pedrycz, W.: Granular Computing with Shadowed Sets. In: Ślęzak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFDGrC 2005. LNCS (LNAI), vol. 3641, pp. 23–32. Springer, Heidelberg (2005)

Pedrycz, W., Gomide, F.: Fuzzy systems engineering: toward human-centric computing. Wiley & Sons, Hoboken (2007)

Peters, J.F., Skowron, A., Synak, P., Ramanna, S.: Rough sets and information granulation. In: De Baets, B., Kaynak, O., Bilgiç, T. (eds.) IFSA 2003. LNCS, vol. 2715, pp. 370–377. Springer, Heidelberg (2003)

Qiu, T., Chen, X., Liu, Q., Hang, H.: A granular space model for ontology learning. In: Proceedings of the IEEE International Conference on Granular Computing, San José, California, pp. 61–65 (2007)

Valente de Oliveira, J.: Semantic Constraints for Membership Function Approximation. IEEE T. Syst. Man Cy. A 29(1), 128–138 (1999)

Zadeh, L.A.: Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Set Syst. 90(2), 111–127 (1997)

Zhong, N., Yao, Y.Y., Qin, Y., Lu, S., Hu, J., Zhou, H.: Towards granular reasoning on the Web. In: Proceedings of the International Workshop on New Forms of Reasoning on the Semantic Web, Bangkok, Thailand (2008)

Yao, J.T.: A ten-year review of granular computing. In: Proceedings of the IEEE International Conference on Granular Computing, San Jose, California, pp. 734–739 (2007)

Yao, J.T., Yao, Y.Y.: Information granulation for web based information retrieval support systems. In: Proceedings of SPIE, Orlando, Florida, pp. 138–146 (2003)

Yao, Y.Y.: Information granulation and rough set approximation. Int. J. Intell. Syst. 16(1), 87–104 (2001)

Yao, Y.Y.: Three perspectives of Granular Computing. J. Nanchang Institute of Technology 25(2), 16–21 (2006)

Yao, Y.Y.: Granular Computing for Web Intelligence and Brain Informatics. In: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, pp. 21–24. Silicon Valley, California (2007)

Yao, Y.Y.: Granular computing: past, present and future. In: Proceedings of the IEEE International Conference on Granular Computing, Hangzhou, China, pp. 80–85 (2008)