Correspondence

Granular Mappings

Andrzej Bargiela and Witold Pedrycz

***Abstract*—In this paper, we are concerned with the granular representation of mappings (or experimental data) coming in the form $R : R \to [0, 1]$ (for one-dimensional cases) and $R : R^n \to [0, 1]$ (for multivariable cases) with R being a set of real numbers. As the name implies, a granular mapping is defined over information granules and maps them into a collection of granules expressed in some output space. The design of the granular mapping is discussed in the case of set and fuzzy set-based granulation. The proposed development is regarded as a two-phase process that comprises: 1) a definition of an interaction between information granules and experimental evidence or existing numeric mapping and 2) the use of these measures of interaction in building an explicit expression for the granular mapping. We show how to develop information granules in case of multidimensional numeric data by resorting to fuzzy clustering (fuzzy C-means). Experimental results serve as an illustration of the proposed approach.**

***Index Terms*—Fuzzy clustering, fuzzy relational equations, granular computing, information granules, possibility and necessity measures, specificity of granular information.**

I. Introduction and Problem Statement

The notion of information granularity and granular computing [1], [15]–[18] has emerged as one of the paradigms of computing that permeates a broad array of pursuits of computational intelligence and contributes as one of its fundamentals. With the notion of information granules and granularity being central to various aspects of problem abstraction and modularization, we encounter different formal ways of expressing information granularity. The notion of granularity is formalized in the language of sets (quite often alluded to interval analysis), fuzzy sets, and rough sets, to name just a few general formal environments [1]–[3], [10], [11]. While information granules (sets, fuzzy sets, rough sets, etc.) are the building entities, the fundamental construct is the one of a mapping. Obviously, when dealing with information granules, the pertinent mapping can be referred to as a granular mapping. In a descriptive manner, by the granular mapping we mean a transformation from some input space to output space that is characterized at the granular level; this means that it operates on information granules defined in the corresponding spaces. Granular mappings can be encountered quite frequently in rule-based systems (e.g., fuzzy rule-based systems), where the mapping is given in a form of "if–then" statements [7]–[9]. We encounter granular mappings in classification problems, cf. [2], [6], [13], [14]. There are two fundamental categories of the problems arising in such setting, that is: 1) analysis of granular mappings (say, rule-based systems), which is inherently associated with various interpretation aspects and 2) design of these mappings which requires the development of the (experimentally meaningful and transparent) associations between the information granules.

More formally, the problem we are interested in is posed as follows:

> Given: 1) a function R or a collection of experimental data (D) defined in some input space $X (X \subset R)$ and assuming values in the unit interval ([0,1]) and 2) a finite collection of information granules $A = \{A_1, A_2, \ldots, A_c\}$ defined in the same input space X; represent R (D) as a granular mapping. That is, construct a transformation $A \to \mathcal{G}(Y)$ with $\mathcal{G}(Y)$ denoting a family of information granules (say, sets, fuzzy sets, rough sets, etc.) defined in the output space $Y (= [0, 1])$.

To assure clarity of presentation, the material is organized in a bottom-up format. First, we elaborate on the concept and realization of the notion of interaction of granular probes with the experimental environment and demonstrate that possibility and necessity measures can be viewed as a viable vehicle that helps quantify such concept. We then move to the simplest one-dimensional (1-D) scenario in which such granular mappings arise and consider a family of sets (intervals) being treated as granular probes in this granular environment. In the sequel, in Section IV, we discuss a multidimensional case in which fuzzy clustering plays an important role.

II. Possibility and Necessity Measure as the Computational Vehicle of Granular Representation

In light of the overall scenario outlined in Section I, we easily envision that a sound starting point pertains to a way in which information granules interact with the mapping (R) or experimental data over which such a granular mapping has to be realized. Intuitively, one of the first realizations that come to mind is the one where we attempt to quantify the interaction between $A_i s$ and the mapping. As the granules adhere to the fundamental concepts of granular mechanisms and computing such as inclusion, intersection, complement, etc. (no matter how these concepts become implemented, which, in turn, depends upon the nature of the specific granular environment), it becomes appealing to revisit two of them. First, a notion of intersection tells us about the interaction between the probe (A_i) and the given mapping. Second, a way in which inclusion of the probes in the mapping expresses another point of view at the interaction. Interestingly, these two ideas have already been in existence in the literature and come under the name of possibility and necessity measures. More specifically, given two information granules A and R (here, A stands for one of the granules from the family of the granular probes A while R is mapping whose granular format we are about to develop), the possibility measure, $\mathrm{Poss}(A, X)$, describes a level of overlap between these two. The necessity measure $\mathrm{Nec}(A, X)$ captures a level of inclusion of A in X. While these descriptors are quite generic, their realization needs to be described in more detail, depending upon the character of the granular environment. In case of fuzzy sets (and sets by this matter), we have [4], [5]

$$\mathrm{Poss}\ (A, R) = \sup_{x \in X} \left[A\ (x) t R(x)\right]. \qquad (1)$$

The plot visualizing the computations of the possibility and necessity measures is shown in Fig. 1. Computationally, we note that the possibility measure looks at the intersection between A and R and then takes an optimistic aggregation of the intersection by picking up the highest

Manuscript received May 28, 2003; revised March 26, 2004. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC), in part by the Alberta Software Engineering Research Consortium (ASERC), and in part by the Engineering and Physical Sciences Research Council (EPSRC-UK). The work of W. Pedrycz was supported in part by the Canada Research Chair (CRC) Program. This paper was recommended by Associate Editor T. Sudkamp.

A. Bargiela is with the School of Computing and Mathematics, Nottingham Trent University, Nottingham NG1 4BU, U.K.

W. Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2G7, Canada, and also with the Systems Research Institute, Polish Academy of Sciences, Warsaw 01-447, Poland.

Digital Object Identifier 10.1109/TSMCA.2005.843381

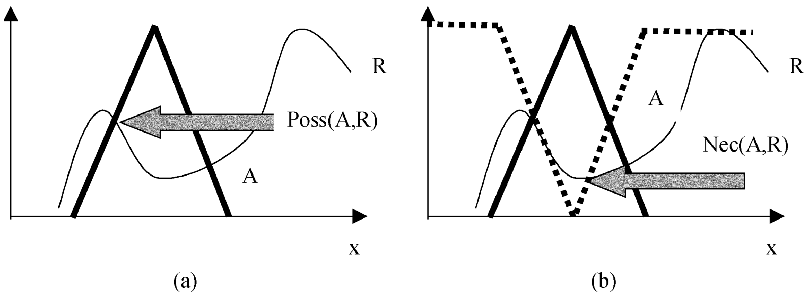

Fig. 1. Computations of (a) possibility and (b) necessity measures treated as a vehicle of interaction between a granular probe A and the numeric mapping R; t-norm: minimum, s-norm: maximum, the dotted line in (b) shows a complement of A, 1-A(x).

values among the intersection grades of A and R that are taken over all elements of the universe of discourse X. The necessity measure expresses a pessimistic degree of inclusion of A in R and is computed as follows, cf. [5]

$$\mathrm{Nec}\ (A, R) = \inf_{x \in X} \left[\Big(1 - A\ (x)\Big) s R(x) \right]. \tag{2}$$

The computational details are revealed in Fig. 1. In contrast to the possibility measure, the necessity measure is asymmetric (which is quite obvious, as we are concerned with the inclusion predicate).

These two definitions are applied in case of R that is given in an analytic (explicit) fashion. If we are provided with the experimental data (namely, input–output pairs $D = \{(x_k, y_k)\}, k = 1, 2, \ldots, N$), then the above calculations are modified with the supremum and infimum operations being replaced by the maximum and minimum operations taken over all data D

$$\mathrm{Poss}\ (A, D) = \max_{(x_k, y_k)} \left[A\ (x_k) \mathrm{t} \mathrm{y}_k \right] \tag{3}$$
$$\mathrm{Nec}\ (A, D) = \min_{(x_k, y_k)} \left[\Big(1 - A\ (x_k)\Big) \mathrm{s} \mathrm{y}_k \right]. \tag{4}$$

The possibility and necessity measures articulate a way in which A interacts with R or experimental data.

Considering the family of the information granules A, we compute the possibility and necessity measures with respect to R or D and end up with the 2c-tuple representation

$$\lambda_i = \mathrm{Poss}(A_i, R) \qquad \mu_i = \mathrm{Nec}(A_i, R) \tag{5}$$

which is a manifestation of R (or D) expressed in the granular language of A. Noticeably, by changing elements of A, we end up with the different representation of the same mapping. Let us reiterate that different types of As provide us with different points of view (perspectives) on the same mapping R or experimental evidence D.

III. Building the Granular Mapping

So far, we have arrived at the representation (manifestation) of R expressed in the languages of $A_i s$. This forms a prerequisite to construct a granular mapping as the elements of A are the basis for the granules in the output space. They are related to $A_i s$ and result through the process of the reconstruction guided by (λ_i, μ_i). From the computational standpoint, we can view this as a solution to a certain inverse problem. Let us start with a single information granule (A) for which the values of λ and μ are known. There is no unique solution to this problem. There is, however, a maximal information granule denoted here by $\hat{R}$, [4], whose construction is supported by the theory of fuzzy relational equations (as a matter of fact, (5) is a $\sup - t$ composition of R and A). The membership (characteristic) function of this maximal fuzzy set (mapping) induced by the A reads as

$$\hat{R}\ (x) = A(x) \to \lambda = \begin{cases} 1, & \text{if } A(x) \leq \lambda \\ \lambda, & \text{otherwise} \end{cases}. \tag{6}$$

The above formula applies to the $t - \mathrm{norm}$ realized as a minimum operator. In general, (6) reads in the form

$$\hat{R}\ (x) = A(x) \to \lambda = \sup \left[a \in [0,1] | \text{at } A(x) \leq \lambda \right]. \tag{7}$$

When using the entire family of $A_i s$ (that leads to the intersection of $\hat{R}_i s$), we obtain

$$\hat{R} = \bigcap_{i=1}^{c} \hat{R}_i. \tag{8}$$

From the theoretical point-of-view that arises in the setting of fuzzy relational equations, we note that we are dealing here with a system of equations $\lambda_i = \mathrm{Poss}(A_i, R), i = 1, 2, \ldots, c$ to be solved with respect to R for λ_i and A_i provided.

The theory of fuzzy relational equations plays the same dominant role in the case of the necessity computations. It is worth noting that we are faced with so-called dual fuzzy relational equations. Here, the minimal solution to (5) for A and μ_i given reads in the form

$$\widetilde{R}(x) = \Big(1 - A\ (x)\Big) \varepsilon \mu = \begin{cases} \mu, & \text{if } 1 - A\ (x) < \mu \\ 0, & \text{otherwise} \end{cases}. \tag{9}$$

Again, the above formula applies to the maximum realization of the $s - \mathrm{norm}$. The general formula takes the form of

$$\tilde{R}(x) = \Big(1 - A\ (x)\Big) \varepsilon \mu = \inf \left[a \in [0,1] | \text{as } \Big(1 - A(x)\Big) \leq \mu \right\}. \tag{10}$$

Because of the minimal solution, the collection of the granular probes leads us to the partial results that are afterwards combined through a union operation

$$\widetilde{R} = \bigcup_{i=1}^{c} \widetilde{R}_i. \tag{11}$$

In conclusion, (7) and (10) become the granular representations of the mapping (R) arising in the context of the collection of information granules A given in advance. The obvious containment relationship holds

$$\widetilde{R} \subseteq R \subseteq \hat{R} \tag{12}$$

where the granularity of the mapping manifests through the two different bounds (lower and upper approximations of R).

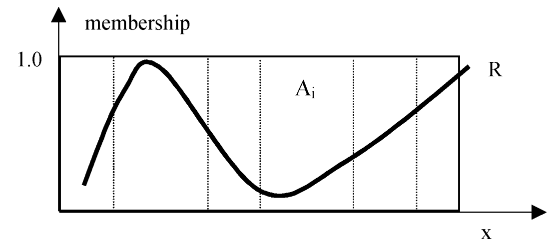

Fig. 2. Mapping (R) with a collection of superimposed sets (intervals) A_i.

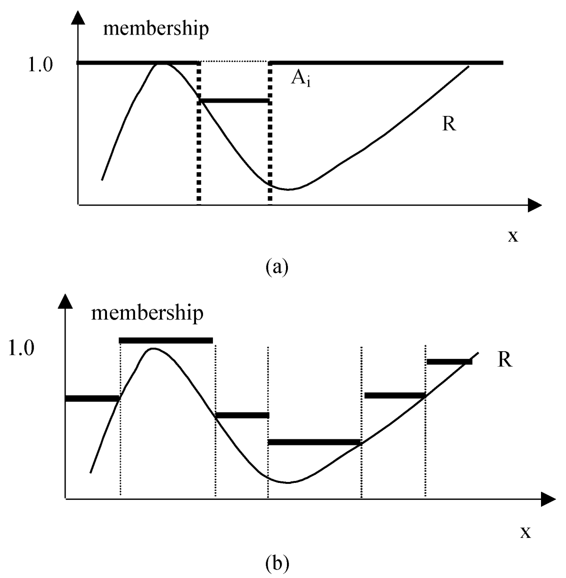

Fig. 3. Computing of the upper bound of R (solid staircase line) with the use of (a) a single set (A) and (b) family $\{A_i\}$.

As a simple, yet highly illustrative example, consider a collection of sets (intervals) regarded as granular probes of some nonlinear numeric mapping (see Fig. 2). A single information granule produces the result shown in Fig. 3(a), which, in fact, produces membership values equal to 1 over any argument not belonging to A. The aggregation of all partial results gives more specific results (see Fig. 3).

Interestingly, the reconstructed fuzzy set exhibits a stairwise type of membership function where the height of the individual jumps and their distribution across the space depends on the distribution of A_is. The same effect that concerns the lower bound of R is present in Fig. 4. When combined together, the result is a granular mapping, Fig. 5. It is worth noting that by changing the position of the cutoff points (intervals), we end up with different granular mappings. Eventually the mapping can be subject to some optimization in which we develop the collection of A_is in such a way that the granular mapping is as specific as possible (so that the bounds are made tight).

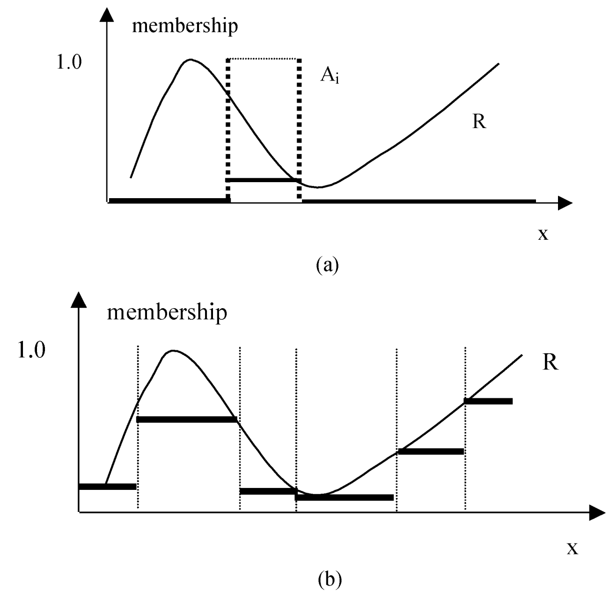

Fig. 4. Mechanism of reconstruction realized in terms of the necessity measure.

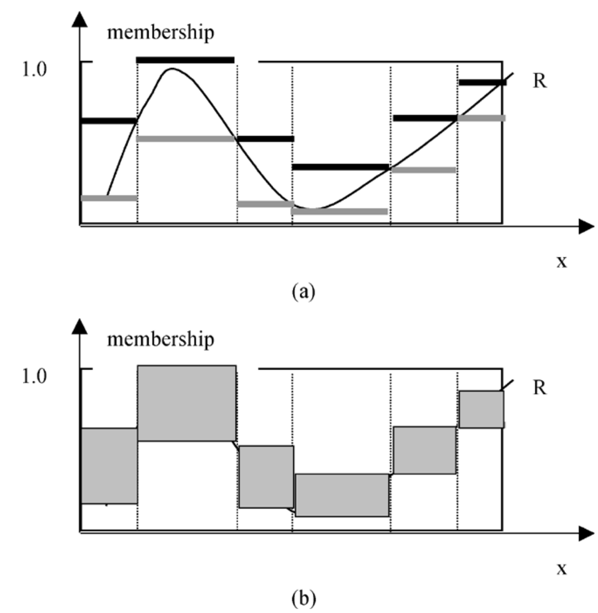

Fig. 5. Building a granular mapping: (a) upper and lower bound of the mapping and (b) interval-valued granules in the output space.

IV. Design of Multivariable Granular Mappings

The 1-D case can be generalized with the same design objective in mind as before. The primary step is to find a collection of information granules in R^n so that they are meaningful constructs (in light of the data being available). The dimensionality of the input space suggests treating all inputs at the same time, rather than discussing each variable separately (which is impractical and leads to a significant number of combinations of such granules). In other words, we focus on forming granular relations (e.g., relations, fuzzy relations, etc.) and this immediately brings us to the idea of fuzzy clustering and its objective function-based realization, such as, e.g., fuzzy C-means (FCM). The FCM is a well-documented clustering mechanism that applies to the data in the input space and produces a fuzzy-partition matrix that, in essence, is a collection of discrete fuzzy relations. As the algorithm is widely reported, for our purposes here, let us recall that in optimizing the given objective function, it goes through a series of iterations in which a partition matrix and prototypes are updated following the two expressions:

1) update of the prototypes

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^m x_k}{\sum_{k=1}^{N} u_{ik}^m} \tag{13}$$

2) update of the partition matrix

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}}} \tag{14}$$

where $U = [u_{ik}]$ is a fuzzy partition ($i = 1, 2, \ldots, c, k = 1, 2, \ldots, N$) and v_i denotes a prototype of the ith cluster. The distance function between the kth data point x_k and the ith prototype is denoted by $\|x_k - v_i\|$ (computationally, it is usually implemented using a weighted Euclidean distance). The fuzzification coefficient ($m > 1$) serves as another adjustable parameter of the algorithm, whose value is selected in advance.

Once the clusters have been determined, the next part of the construction is the same as in the 1-D case. To draw the linkages, let us note that the partition matrix can be represented as c fuzzy relations $A_1, A_2, \ldots, A_c$ defined in the finite data set that is

$$U = \begin{bmatrix} A_1 \\ A_2 \\ \\ A_c \end{bmatrix}.$$

By doing this, we explicitly form the collection of the information granules A as was done in the 1-D case. More specifically, the possibility and necessity computed for the granules are governed by the expressions $\mathrm{Poss}\ (A_i, D) = \max_{x_k}[A_i(x_k) t y_k]$ and $\mathrm{Nec}\ (A_i, D) = \min_{x_k}[(1 - A_i(x_k)) s y_k]$. Alluding to the way in which information granules interact with the mapping or data, we note that most of the interaction occurs in the region where the granules are normal (viz. they assume the membership grade equal to 1.0). As the FCM imposes the unity constraint (viz., the membership grades in all clusters assumed for the same point sum up to 1.0), we normalize these grades by elevating the highest membership grade to 1.0.

V. Role of the Design Parameters in the Construction of Granular Mappings

There are two essential parameters of the granular mapping that result from the use of the fuzzy clustering in the development of the underlying construct. These are the number of the clusters (c) and the shape of the probes and possible patterns of behavior (interaction) between them.

We have already seen that the number of information granules dictates the granularity of the mapping. This has been clearly demonstrated in the case of sets (intervals) in the 1-D problem in which a certain portion of the data (or function) fall under the realm of the granule (become identified by it) and then imply the granularity of the mapping itself. We have learned that in limit (which may not be of practical interest), the granularity of the mapping is very high. The quantification of this effect could be put forward by computing a ratio of the average value of σ-count (cardinality) of the information granules in the output space to the granularity of the probes (fuzzy relations in the input space).

As the clusters regarded in the context of the design of the granular mapping play a primordial role (as our intuition might have already suggested), it is instructive to get a better grasp as to our possible control over the shape of the clusters. The illustration in a 1-D case is the best option. Fig. 6 shows the membership functions for selected values of the fuzzification coefficient (the values of the prototypes are kept fixed and equal to 1–3, respectively). Depending upon its value of this coefficient, the resulting fuzzy sets tend to resemble sets (when m approaches 1) or exhibit a significant overlap and interact between themselves (which is quite visible for the values of m around 2) or get quite

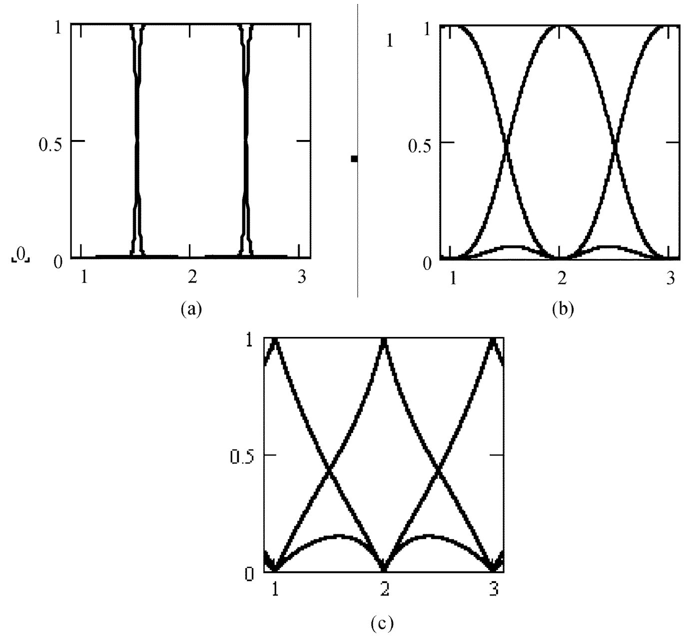

Fig. 6. Plots of membership functions of information granules generated by the FCM clustering for selected values of m: (a) $m = 1.1$, (b) $m = 2.0$, and (c) $m = 3.0$.

"spiky" and loose most of the interaction with the increasing values of the fuzzification coefficient (a good example is $m = 3$).

VI. Quantification of Granular Mappings

As underlined before, each segment of the granular mapping is formed within a scope of the individual information granule occurring in the input space. The granularity of the mapping (quantified at the end of the output space) is directly related with the upper and lower bound of the interval formed by the possibility and necessity values. To quantify this granularity and relate it to the experimental data (which is the entry point of the overall design), it is beneficial to introduce the following index:

$$\overline{g} = \frac{1}{N} \sum_{k=1}^{N} (u_k - l_k) \tag{15}$$

which tells us about an average spread of the interval with l_k and u_k denoting the lower and upper bound, respectively, being computed as given (6) and (9). The average over N data points helps us quantify the granulation effect and abstract it from the individual variations. Higher values of g imply lower granularity of the resulting mapping (that is, larger differences between the bounds of the mapping).

VII. Experimental Studies

The experiments reported here aim at visualizing the performance of the granular mappings in case of numeric data. The data set comes from the machine learning repository and deals with fuel economy (in miles per gallon) of various vehicles described by a series of seven parameters including displacement, weight, number of cylinders, etc. The output has been normalized to the unit interval. Furthermore, the dataset was split randomly into 50%–50% training and testing subset. The training set was used to carry out clustering, and then we computed the possibility–necessity characteristics of the granular mapping. The FCM was run for 20 iterations (at which point there were practically no visible changes in the values of the objective function). The distance function

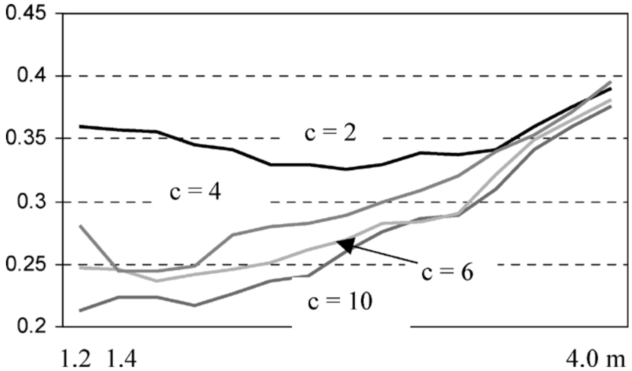

Fig. 7. Performance (granulation) index of the mapping as a function of the fuzzification coefficient of the FCM for selected number of the clusters (c).

TABLE I
PERFORMANCE OF THE GRANULAR MAPPING ON A TRAINING AND TESTING SET

c	2	4	6	10
Training set	0.326	0.245	0.237	0.213
Testing set	0.354	0.256	0.256	0.212

was assumed to be the normalized Euclidean (the weights were computed as the standard deviations of the individual inputs).

There are two essential parameters of the mapping associated with the clustering mechanism that is the number of clusters (information granules to realize the mapping) and the fuzzification coefficient (m). We explore various combinations of these two to obtain a general sense as to their role and some general tendencies. The main results are displayed in Fig. 7. The figure helps us to draw several conclusions. First, as one could have intuitively expected, the increased number of clusters yields better performance of the mapping (the bounds start to be tighter). This tendency becomes quite apparent; the changes in the performance become more evident for the lower number of clusters (hence, there is a significant difference when moving from two to six clusters and far less accentuated difference for higher values of c). The number of clusters also affects the optimal values of the fuzzification coefficient. In general, we witness a tendency of lower optimal values of m with the increased number of the clusters. Relating this to the shape of the membership functions (see Fig. 6), one could envision that with more clusters their boundaries need to be shaped as more set-oriented and with less spread and overlap with the neighboring information granules. More specifically, the optimal fuzzification factors are given as: $c = 2$, $m = 2.6$; $c = 4$, $m = 1.6$; $c = 6$, $m = 1.6$; $c = 10$, and $m = 1.2$.

The testing set was used to assess the generalization abilities of the granular mapping. The results shown in Table I indicate that the differences in performance on the training and testing set are not substantial which leads us to conclude about the sound generalization abilities. In essence with $c = 10$, the performance on the training and testing set is practically the same.

Figs. 8 and 9 visualize the behavior of the granular mapping for selected number of clusters (with the optimal values of the fuzzification coefficients). In these two cases, where $c = 2$ and $c = 10$, the essential component of the granular mapping—possibility and necessity measures are included in Table II.

We can present the distribution of bounds by plotting individual differences between the bounds (lower and upper) and the numeric experimental data, see Fig. 10.

Noticeably, the bounds are asymmetric; the two plots above help us in two ways: 1) they are useful in identifying data points with the broadest intervals that could be revisited as potential outliers and 2) we learn about the distribution of granularity of the realized mapping and

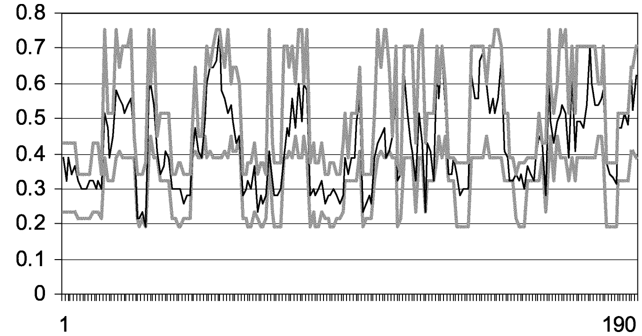

Fig. 8. Training data and bounds of the granular mapping ($c = 2$).

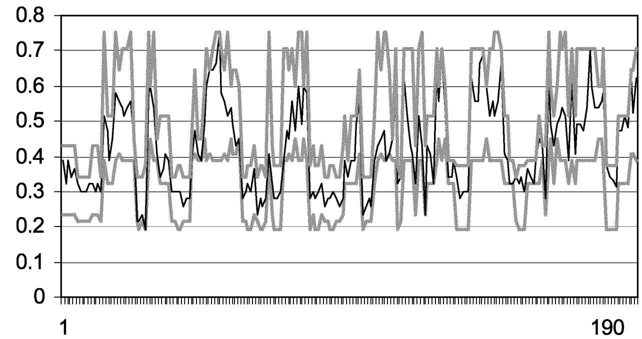

Fig. 9. Training data and bounds of the granular mapping ($c = 10$).

TABLE II
POSSIBILITY AND NECESSITY MEASURES OF GRANULAR MAPPING: (a) $c = 2$ AND (b) $c = 10$. IN BOTH CASES THE RESULTS ARE REPORTED FOR THE OPTIMAL VALUES OF THE FUZZIFICATION FACTOR

Cluster no.	Possibility	Necessity
1	0.483	0.193
2	0.751	0.321

(a)

Cluster no.	Possibility	Necessity
1	0.601	0.451
2	0.644	0.408
3	0.708	0.386
4	0.386	0.322
5	0.429	0.236
6	0.451	0.386
7	0.376	0.193
8	0.751	0.386
9	0.515	0.322
10	0.343	0.215

(b)

quantify which bound (lower or upper) is "tighter," with respect to the experimental data.

VIII. CONCLUDING COMMENTS

We have introduced and studied the concept, analyzed properties, and discussed the design of granular mappings being viewed as one of the fundamental constructs of granular computing. The developed granular mapping helps establish a global and more general view on data or detailed numeric functions that, for this purpose, are perceived as a collection of information granules. In a nutshell, the introduced construct is formed by probing the function or data by a series of granular "probes" (sets or fuzzy sets), recording the results of such an interaction, and aggregating the results. The underlying logical framework

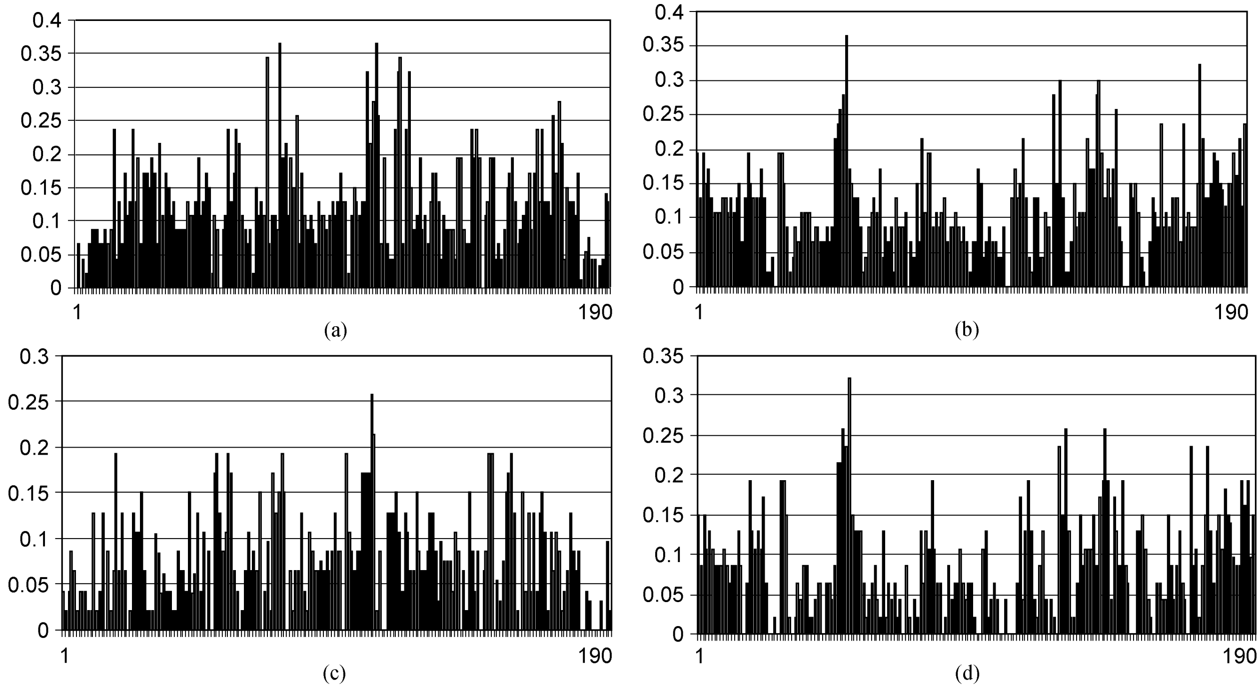

Fig. 10. Differences between the bounds and numeric experimental data for $c = 4$, $m = 1.4$: (a) upper bound and data; (b) lower bound and data $c = 18$ $(m = 1.2)$; (c) upper bound and data; and (d) lower bound and data.

is based on the calculus of fuzzy relational equations. The probing results come in the form of possibility and necessity measures. In the sequel, these give rise to a set of fuzzy relational equations and the development of the granular mapping is formed by solving a system of such equations. The resulting lower and upper bounds fully describe the mapping. We elaborate on the role and optimization of the granular probes and show how their selection is realized through fuzzy clustering (and FCM, in particular). The choice of the shape of the fuzzy relations (which is controlled by the values of the fuzzification coefficient) and the number of clusters (granules) is studied in detail with a quantitative demonstration of their impact on the specificity of the granular mapping.

References

[1] A. Bargiela and W. Pedrycz, *Granular Computing: An Introduction*. Norwell, MA: Kluwer, 2002.

[2] ——, "Recursive information granulation: aggregation and interpretation issues," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 33, no. 1, pp. 96–112, Jan. 2003.

[3] G. Bortolan and W. Pedrycz, "Reconstruction problem and information granularity," *IEEE Trans. Fuzzy Systems*, vol. 5, no. 2, pp. 234–248, May 1997.

[4] A. Di Nola, W. Pedrycz, S. Sessa, and E. Sanchez, *Fuzzy Relational Equations and Their Applications in Knowledge Engineering*. Dordrecht, Germany: Kluwer, 1989.

[5] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York: Academic, 1980.

[6] B. Gabrys and A. Bargiela, "General fuzzy min-max neural network for clustering and classification," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 769–783, May 2000.

[7] T. Y. Lin, "Neural networks, qualitative fuzzy logic and granular adaptive systems," in *Proc. World Congress Computat. Intell.*, Honolulu, HI, May 12–17, 2002, pp. 566–571.

[8] ——, "Rough logic formalism for fuzzy controllers—a hard and soft computing approach," *J. Approx. Reasoning*, vol. 15, pp. 95–414, 1996.

[9] ——, "Data mining: granular computing approach," in *Lecture Notes Artif. Intell. 1574, 3rd Pacific Asia Conf.*, Beijing, China, Apr. 26–29, 1999, pp. 24–33.

[10] R. Moore, *Interval Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1966.

[11] S. K. Pal and A. Skowron, Eds., *Rough Fuzzy Hybridization. A New trend in Decision-Making*, Singapore: Springer Verlag, 1999.

[12] W. Pedrycz and F. Gomide, *An Introduction to Fuzzy Sets: Analysis and Design*. Cambridge, MA: MIT Press, 1998.

[13] W. Pedrycz and A. Bargiela, "Granular clustering: a granular signature of data," *IEEE Trans. Systems, Man, Cybern.* , vol. 32, no. 2, pp. 212–224, Feb. 2002.

[14] P. K. Simpson, "Min-max neural networks—Part I: classification," *IEEE Trans. Neural Netw.*, vol. 3, no. 5, pp. 776–786, May 1992.

[15] L. A. Zadeh, "Fuzzy sets and information granularity," in *Advances in Fuzzy Set Theory and Applications*, M. M. Gupta, R. K. Ragade, and R. R. Yager, Eds. Amsterdam, The Netherlands: North Holland, 1979, pp. 3–18.

[16] ——, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets Syst.*, vol. 90, pp. 111–117, 1997.

[17] ——, "Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems," *Soft Comput.*, vol. 2, pp. 23–25, 1998.

[18] L. A. Zadeh and J. Kacprzyk, *Computing With Words in Information/Intelligent Systems*. Heidelberg, Germany: Physica-Verlag, 1999, vol. 1–2.