Int. J. Appl. Math. and Comp. Sci., 1999, Vol.9, No.2, 435–455

ARCHITECTURES OF GRANULAR INFORMATION AND THEIR ROBUSTNESS PROPERTIES: A SHADOWED SETS APPROACH

WITOLD PEDRYCZ*

This paper addresses an important issue of information of granulation and relationships between the size of information granules and the ensuing robustness aspects. The use of shadowed sets helps identify and quantify absorption properties of set-based information granules. Discussed is also a problem of determining an optimal level of information granulation arising in the presence of noisy data. The study proposes a new architecture of granular computing involving continuous and granulated variables. Numerical examples are also included.

Keywords: information granulation and information granules, fuzzy sets, shadowed sets, noisy data, uncertainty, generalized multiplexer, switching mechanisms.

1. Introduction

The processes of information granulation and the resulting information granules become a cornerstone of most human pursuits. We encapsulate a lot (sometimes an infinite number) of elements into a single entity, label it, attach a well-defined and useful meaning and start using it in the representation of the problem at hand. Information granules are also building blocks of any algorithms. The specificity (granularity) of such granules helps combat or reduce complexity of the overall problem through its suitable modularization. There are several main avenues supporting the development of information granules

- Set theory, especially interval analysis, cf. (Milanese *et al.*, 1996; Moore, 1966; 1979; 1988). Interval analysis comes as one of the earliest manifestation of information granulation that hinges on intervals (sets in the space of reals $\mathbb{R}$). With its roots originating from numerical and tolerance analysis, interval analysis regards intervals as crucial information granules involved in any computer computations. An important stream of investigation in this realm concerns control problems where intervals and interval analysis emerge as a consequence of set-theoretic view at system's disturbances (giving rise to a model of so-called bounded and unknown disturbances) and limited accuracy of model estimation (identification) procedures.

* Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2G7 Canada, e-mail: pedrycz@ee.ualberta.ca.

- Fuzzy sets (Kandel, 1986; Pedrycz, 1996; 1997; 1998; Pedrycz and Gomide, 1998; Zadeh, 1979) admit continuos boundaries between the notion of exclusion and inclusion of elements and arise as a plausible model of describing linguistic concepts of ill-defined boundaries where the fundamental set-based dichotomy becomes irrelevant and highly artificial. The concept of continuous and gradual changes in membership values becomes an attractive alternative in coping with an array of situations involving human experts and formalizing their qualitative domain knowledge. Fuzzy sets come also with their interesting generalizations: fuzzy sets of higher order and higher types, interval-valued fuzzy sets (Sambuc, 1975), L-fuzzy sets, etc.
- Rough sets developed in (Pawlak, 1982; 1991), see also (Skowron, 1990; Slowinski, 1992) exhibit another position as to the construction of information granules by assuming an existence of some indiscernability relation that becomes a cornerstone of any resulting entities.

The selection of one of the formal frameworks implies a full commitment to all their features as well as a strong confinement to their eventual limitations being inherently associated with the formalism of operations on information granules available therein.

There is a strong motivation to granulate information, no matter which way one decides to proceed:

- Information granules help organize and perceive the problem at hand in a more organized and clear fashion. Rather than viewing the problem in terms of a vast array of numeric data, we perceive it in the framework of far more meaningful information granules such as intervals or linguistic terms (fuzzy sets). The paradigm of rule-based architectures is an excellent example of an extensive use of information granules. Information granules occur in each rule. Each rule comes as a piece of 'local' operational domain knowledge.
- Information granules allow us to make problems computationally manageable. As the resulting computing architectures are highly modular (see e.g., rule-based systems), this facet promotes a partition of the original problem and reduce an overall computing effort.

While information granules, information granulation, and granular computing, in general, constitute a useful concept and powerful computing paradigm, there are still a number of open questions and essential research pursuits to be prudently investigated. What becomes a particularly burning issue is a determination of the size of the information granules arising in a particular problem at hand. Information granules interact with an external world. The granularity and character of such entities are predominantly affected by the environment of the problem. One needs to be cognizant as to the consequences of the size of the information granules and their distribution in the individual spaces.

This study is geared toward this important problem. In order to make the discussion more focused, we concentrate on the following environment. We consider information granules being modeled as sets. The computing environment formed by them interacts with data that are noisy. The objective is to study the form of the resulting interaction, analyze a way in which it can cope with randomness of the data, and reveal and quantify crucial relationships between the granularity of the computing environment and probabilistic characteristics of the data.

The material is organized as follows: First, we discuss shadowed sets to be viewed in this study as a useful vehicle conveying the aspects of robustness of set-based information granules. Afterwards, in Section 2, we define the notion of robustness itself and revisit the issue of noisy data in the set-theoretic setting. This leads us to its further quantification of uncertainty stemming from noisy data in the form of shadowed sets of the original set-based information granules. A detailed description of the absorption regions is discussed in Sections 4 and 5; in the first case we elaborate on the use of the shadowed sets; in the latter concentrate on the induced fuzzy sets. Section 6 is devoted to the problem modularization with the use of granular information and its links with the underlying absorption mechanism. A generalized architecture of a multiplexer along with detailed examples is discussed in Section 7. Concluding remarks are covered in Section 8.

2. Shadowed Sets: Their Rationale and Main Properties

The idea of shadowed sets was originally introduced in (Pedrycz, 1998) with an intent of breaking out of a long-lasting and quite restrictive tradition of describing information granules via precise, purely numeric membership values. Let us begin with a formal definition of this structure. A shadowed set A defined in $\boldsymbol{X}$ is characterized by a three-valued membership function

$$A:\ \boldsymbol{X} \to \{0, [0,1], 1\}$$

Note that all elements of $\boldsymbol{X}$ being mapped by A to 0 are those completely excluded from the shadowed set. All elements of $\boldsymbol{X}$ that A maps to 1 fully belong to the shadowed set. Finally, there are some elements of $\boldsymbol{X}$ with no specific numeric grade of membership assigned. We associate with them the entire unit interval. This character of membership assignment underlines a fact that *any* membership value is equally possible or preferred. This multivalued (relational) assignment makes the shadow completely unspecified without necessarily forcing us to confine ourselves to any particular numeric membership value. The introduced construct of shadowed sets moves beyond the boundaries of purely numeric membership values. An example of A is visualized in Fig. 1. Observe the shadowed regions of the membership values that apparently led to the name of this concept. On the other hand, if these regions are reduced to zero, then we end up with the generic form of a set.

Let us establish the following notation: the core (elements of $\boldsymbol{X}$) fully belonging to the shadowed set A will be denoted by Core(A). The left- and right-hand shadow of A will be denoted by $\mathrm{Shad}_-(A)$ and $\mathrm{Shad}_+(A)$, respectively.

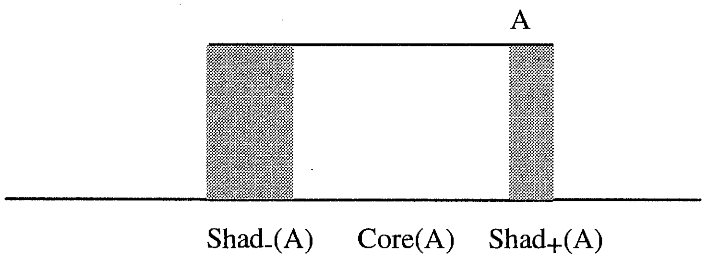

Fig. 1. An example of a shadowed set.

Shadowed sets, as introduced and studied in (Pedrycz, 1998), were conceived from the standpoint of fuzzy sets. As a matter of fact, they are essentially induced by fuzzy sets. The motivation behind such construction was to 'localize' uncertainty and the concept of partial membership conveyed by any fuzzy set and provide with a simplified vehicle for computing with fuzzy sets. As will become apparent in this study, the development of the shadowed sets here is motivated by the existing factor of uncertainty conveyed by data themselves. Additionally, shadowed sets deliver an evident computational advantage over very detailed and quite computationally demanding fuzzy sets.

For the sake of completeness, let us recall the main operations (union, intersection, and complement) defined for the shadowed sets. They are summarized below in the tabular format; here the term *shadow* corresponds to the unit interval:

$A\backslash B$	0	shadow	1
0	0	shadow	1
shadow	shadow	shadow	1
1	1	1	1

$A \cup B$

$A\backslash B$	0	shadow	1
0	0	0	0
shadow	0	shadow	shadow
1	0	shadow	1

$A \cap B$

A	complement
0	1
shadow	shadow
1	0

It is worth underlining that during the operations on shadowed sets, there is a primordial effect of degradation of quality of information meaning that the result becomes a shadow rather than retains an original single numeric membership value. These cases of shadows arising through the logic aggregation of the entries of the shadowed sets are clearly visualized in the above tables.

Shadowed sets are one-dimensional constructs. In other words, they deal with a single variable that is quantified and captured by shadowed sets. There is a straightforward generalization arising in the form of the shadowed relations. Formally speaking, a shadowed relation R defined in a Cartesian product of p universes of discourse (coordinates), say $\boldsymbol{X} = \times_{i=1}^{p} \boldsymbol{X}_i$ is defined by the following set-valued mapping:

$$R: \quad \mathop{\times}_{i=1}^{p} \boldsymbol{X}_i \to \{0, [0,1], 1\}$$

The findings of the following discussion (even though we confine ourselves to shadowed sets) naturally extend to the multidimensional (viz. relation-driven) scenarios.

It is advantageous to underline the main conceptual difference between shadowed sets and rough sets. Making this distinction evident helps us also gain a better insight into the very nature of the ensuing methodological frameworks.

Rough sets (Lin and Wildberger, 1994; Lin and Cercone, 1997) are geared into the analysis of information systems whose attributes (variables) assume a finite number of values. Not moving into details, let us concentrate on a simple example shown in Fig. 2 that exemplifies the essence of rough sets. The universe of discourse (attribute) $\boldsymbol{X}$ has a number of values denoted here by $x_1, x_2, \ldots, x_n$. Now let X be a set. We express it in terms of the values of the attributes. Evidently, this process is not unique. We express X through the lower and upper approximation of X, that is values of $\boldsymbol{X}$ that are fully included in X (lower approximation) and those whose intersection with X is nonempty (upper approximation). The boundary region is defined as a difference (in set-theoretic sense) between the upper and lower approximation.

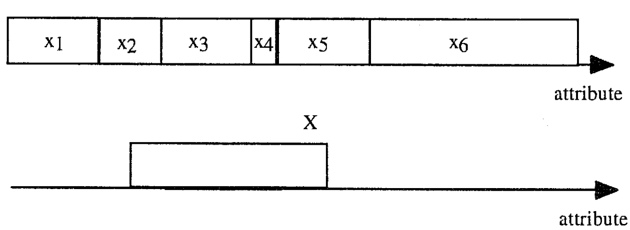

Fig. 2. A set expressed in terms of the values of the attribute: a formation of its lower and upper approximation.

In the situation displayed in Fig. 2, the lower bound of X, denoted by X_*, is equal to $X_* = \{x_3, x_4\}$. The upper approximation, X^*, reads as $X^* = \{x_2, x_3, x_4, x_5\}$. Finally, the boundary region δX is equal to $\{x_2, x_5\}$.

3. Robustness Aspects of Set-Based Information Granules

Sets promote robustness. This phenomenon is invoked owing to the 'absorption' aspects of sets. To clarify the term, let us consider the situation portrayed in Fig. 3.

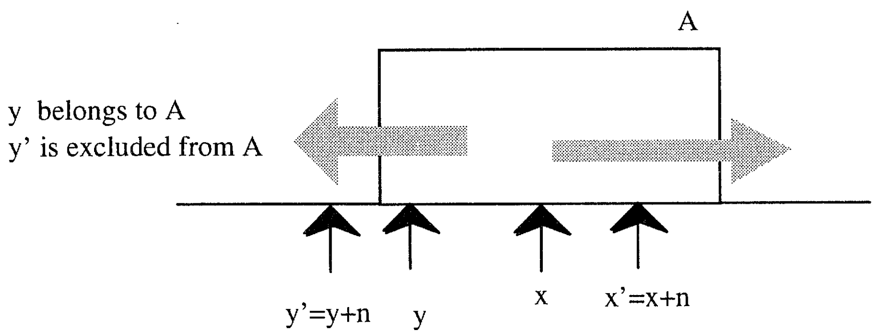

Fig. 3. An absorption effect realized by sets; a noisy input is still encoded without error by 'activating' the right information granule (set) (still being localized within this set).

An original numeric input datum x invokes one of the information granules forming a partition of the space. In other words, we obtain a mapping from numeric values to the set of indexes of the sets. The absorption helps ignore noisy inputs. One may have the original input datum (x) corrupted by noise and resulting as another numeric entry, say $x+n$. This new entity, when confronted with the same information granule (set), gives rise to the proper encoding (viz. it invokes the same set). This is not surprising as the absorption effect and an ability to ignore noisy data is the key rationale behind the introduction of sets and their further usage. Obviously, the main drawback is that sets are of far lower information specificity (lower granularity) in comparison with the original numeric data.

As shown in Fig. 3, there are some limits to the absorption effect conveyed by sets. When a datum (and subsequently, its noisy version) is located quite close to the boundaries of the set, the noise can lead to completely improper encoding. The chances of running into this lack of absorption depend on the level of noise and the position of the data point with respect to the set under consideration (or, indirectly, the granularity of A itself). Our intent is to quantify this effect and propose some protection or awareness mechanisms to be activated in such borderline cases. We reveal that shadowed sets are just an ideal vehicle supporting this development. Let us start with a closer look at the noisy data and propose their set-based interpretation. Subsequently, we use these sets in the induction of the shadowed sets.

4. Noisy Data and Their Set-Based Representation

The basic model of noisy data usually arises in an additive form. Let x denote an original datum whereas its noisy version x' reads as

$$x' = x + n$$

where n is a noise component modeled as a random variable. As usual, we assume that the noise is described by some probability density function (pdf) with zero mean and some standard deviation. Because of the randomness coming into the picture, x' is also a random variable with the same pdf as the original noise and the mean value equal to x. As the already discussed information granules are represented as sets, we would like to maintain the same set-based terminology and algorithmic setting with regard to the noisy data. To accomplish this, it is worthwhile to place x' in the same set-oriented framework. The most appealing way of doing that is through the confidence intervals implied by the pdf. Denote by X a set distributed symmetrically around x whose bounds result from the relationship

$$\int_{x-D}^{x+D} p(y)\,\mathrm{d}y = 1 - \beta$$

Here β, $\beta \in [0,1]$ plays the role of the confidence level. The confidence interval exhibits a straightforward probabilistic interpretation: with probability $1-\beta$ we have the datum embraced by this interval. Evidently, lower values of β produce broader sets, that are information granules of lower specificity (lower granularity). The width of X equal to $2D$ depends on the pdf itself as well as the values of their parameters (in particular, its variance).

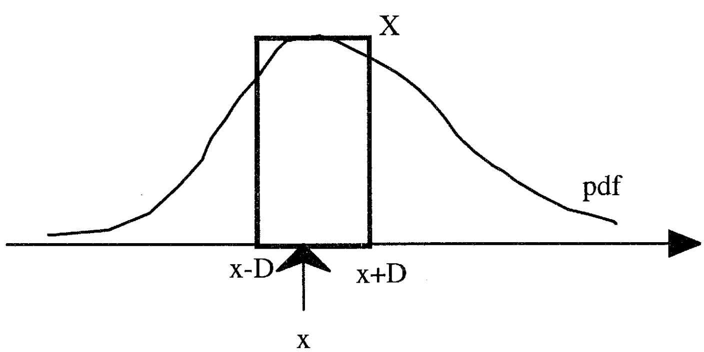

Fig. 4. Computing a confidence set based on the numeric datum and its associated pdf.

Let us provide some illustrative examples for several selected forms of the probabilistic characteristics of the noise coming with the datum. The one commonly encountered in practice is a Gaussian type of noise. In this case, the relationship between the length of the confidence interval (D) and the standard deviation of the Gaussian noise for several values of its standard deviation is shown in Fig. 5.

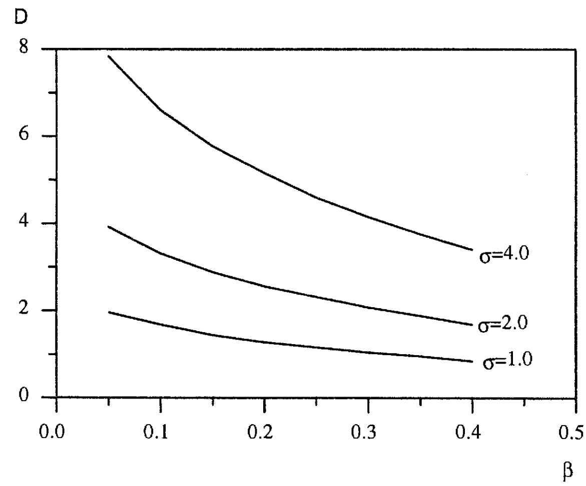

Fig. 5. D viewed as a function of the confidence level β for selected values of the standard deviation of the Gaussian noise.

Similarly as before, we derive the relationships for the confidence interval arising in the presence of the exponential noise with the pdf equal to

$$p(x) = \frac{1}{a}\exp(-x/a)$$

Let us now consider a uniform pdf defined in $[-a, a]$, namely $p(x) = 1/2a$ over this interval and zero otherwise. The calculations of the confidence interval result from the obvious relationship

$$\int_{0}^{x+D} p(y)\,\mathrm{d}y = (1-\beta)/2$$

that gives rise to the linear relationship between b and D,

$$D = a(1-\beta)$$

The highest confidence, $1-\beta = 1$, yields the entire support of the pdf, $D = a$.

Thus, once agreeing on the value of confidence level β, we construct the set-based representation of the noisy datum. This follows the scheme

$$X' \xrightarrow{\beta} [x_-, x_+]$$

This model enters the next step of the design during which we characterize absorption regions of set-based information granules.

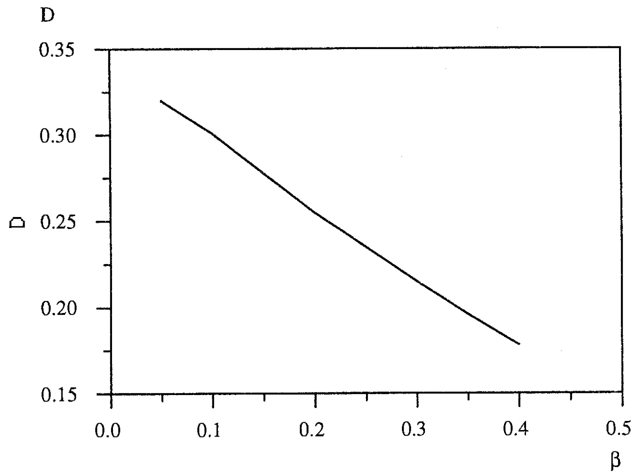

Fig. 6. D viewed as a function of confidence β for selected levels of standard deviation of the exponential noise; $a = 0.5$.

5. Description of Absorption Regions

In what follows, we establish a way in which X' interacts with the sets of the partition $\mathcal{A}$ of the universe of discourse $\boldsymbol{X}$ and identify regions where we start loosing the absorption property of the set-based information granules.

Referring to Fig. 7, it becomes apparent that when X' moves toward the boundaries of a certain set (A_i), then we reach a situation where A_i cannot longer absorb X'; this occurs in the case shown in Figure 7(b). This implies that some regions (of width D) situated at the boundaries of A_i should be viewed differently as the rest of the set. At the location identified there, we regard them as 'uncertain' implying that we cannot view them with confidence as to the absorption abilities. If so, this gives rise to the notion of the shadowed sets: the noisy inputs imply that the original information granule A_i becomes interpreted as the shadow set with the size of the shadow determined by the noise intensity (or, equivalently, the granularity of X'). The same description applies to the remaining sets of the partition, leaving us with the collection of the shadowed sets, cf. Fig. 7(c).

Interestingly enough, the obtained shadowed sets are completely induced by the characteristics of the noise and its intensity. The higher the intensity (standard deviation) of the noise, the more extensive the shadows of the ensuing shadowed sets. In this way, we have arrived at a complete algorithm implying the structure of the shadowed sets. Note that in (Sambuc, 1975) another alternative was discussed using which shadowed sets were induced by the already available fuzzy sets.

There are some straightforward, yet useful insights into the issue of information granularity and the set-oriented characteristics of the noisy data:

- The higher the intensity of noise, the broader the shadows of the induced information granules. The detailed relationship depends on the type of noise.

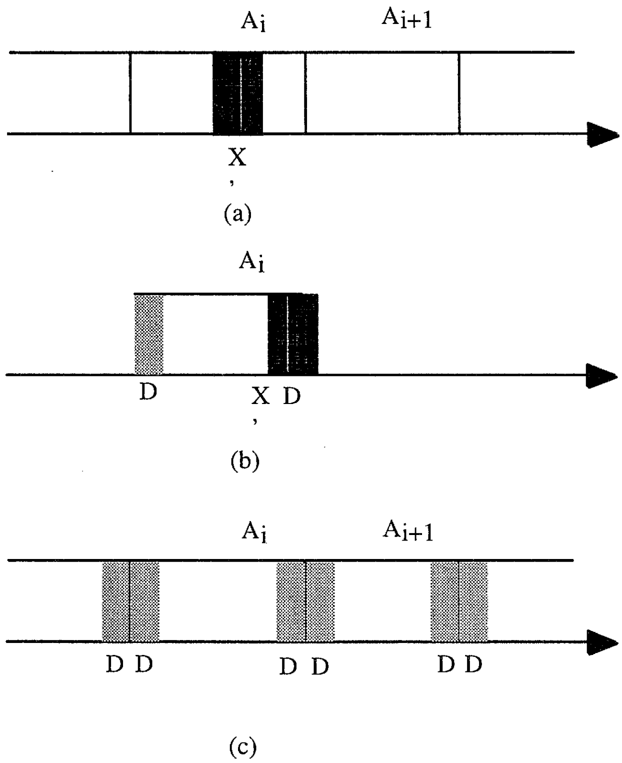

Fig. 7. A set-based absorption effect: X' confined to A_i (a), X' too close to the bounds of A_i and not confined by this information granule (b), conversion of A_i's into their shadow set equivalents.

- The granularity of the partition of the space depends on the intensity of noise we are faced with. A sound design criterion predetermining the size of the information granules can be expressed as follows: for the given noise level, the set-based information granules should be designed in such a way that the induced shadowed sets resulting there do not exhibit empty cores. The underlying point is that when faced with highly noisy information, a too fine partition of $\boldsymbol{X}$ (that involves too specific sets) does not make sense. It is an important design criterion to adhere to. These guidelines strongly emphasize the need of not being overly detailed, if not necessary.

- So far we have assumed that the characteristics of noise are independent of the position x at $\boldsymbol{X}$. It could well be that the intensity of noise is higher at the ends of $\boldsymbol{X}$ while assumes lower values in the middle of the range. If so, this easily leads to the construction of a nonuniform partition of $\boldsymbol{X}$ with more specific sets A_i located in the center of $\boldsymbol{X}$.

6. Fuzzy-Set Based Information Granulation of X

Once we have constructed shadowed sets, they can be further used to construct fuzzy sets. This is an interesting avenue when we start with noisy data and try to determine fuzzy sets that are 'consistent' with them. The crux of this construct is to concentrate on the obtained shadows as those are the regions with the dominant uncertainty factor. What then is proposed is eventually the simplest model of membership transition in the form of the linear membership function as illustrated in Fig. 8. One may think of some other forms of the fuzzy sets, however there is no evidence of their superiority over the simple ones.

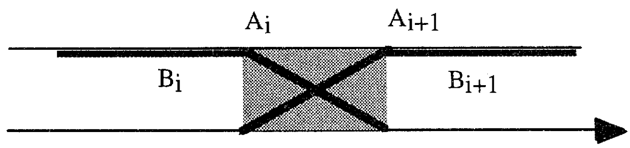

Fig. 8. Constructing fuzzy sets (B_is) with the aid of shadowed sets (A_i's); note that intermediate membership values become associated with the shadows of the shadowed sets.

7. Problem Modularization via Information Granulation

The granulation of the discussed universe of discourse X carried out in terms of A_i's usually done as one of the steps to modularize a problem at hand. In other words, we may think of each A_i as associated with some local algorithm T_i producing a certain outcome (control action, decision, etc.), namely $T_i(x; u, w, \dots, z, \boldsymbol{p})$ where $u, w, \dots, z$ are some other variables occurring in the problem while $\boldsymbol{p}$ denotes the associated vector of the parameters of the algorithm. The partition of X into nonoverlapping sets gives rise to the overall description assuming the following form:

$$y = \begin{cases} T_1 \text{ if } x \in A_1 \\ T_2 \text{ if } x \in A_2 \\ \vdots \\ T_n \text{ if } x \in A_n \end{cases}$$

where the switching between the algorithms is completely determined by the distribution of the individual components of $\boldsymbol{A}$. It is worth noting that when x is affected by noise, so are the results of the algorithm (more precisely, this may result in an invocation of the improper algorithm).

Let us investigate essential implications stemming from the introduction of the shadowed sets. In a nutshell, the resulting shadows produce some 'buffer' (or safety) zones as the local algorithm becomes involved only if x falls in the core of the respective shadowed set. In other words, we get

$$y = \begin{cases} T_1 \text{ if } x \in \text{core}(A_1) \\ T_2 \text{ if } x \in \text{core}(A_2) \\ \vdots \\ T_n \text{ if } x \in \text{core}(A_n) \end{cases}$$

The above solution is obviously incomplete as the shadows are left out and there are no specific values of the control action to be applied at this region. This shortcoming needs to be alleviated. Two conceptually different fixes are considered:

- An explicit expression of the uncertainty factor. In this form of completing the switching between the algorithms, we explicitly admit that there is some hesitation between the choice of the algorithm. Instead of going for one of them (and committing eventually a gross error), our position is to proceed prudently and make use of the two. This leads to the expression

$$y = \begin{cases} T_1 & \text{if } x \in \text{core}(A_1) \\ [T_1, T_2] & \text{if } x \in \text{Shad}_+(A_1) \cup \text{Shad}_-(A_2) \\ & \vdots \\ T_n & \text{if } x \in \text{core}(A_n) \end{cases}$$

Note that in the regions of the shadows, we come up with a relational effect (or-like combination) that fully reflects the uncertainty arising therein. The activation of such aggregation can serve as an important flagging mechanism pointing out at the issue of nonuniqueness of the action.

- An implicit expression of the factor of uncertainty manifesting in some form of aggregation 'agg' of actions of the two pertinent algorithms. This results in the following expression:

$$y = \begin{cases} T_1 & \text{if } x \in \text{core}(A_1) \\ \text{agg}(T_1, T_2) & \text{if } x \in \text{Shad}_+(A_1) \cup \text{Shad}_-(A_2) \\ & \vdots \\ T_n & \text{if } x \in \text{core}(A_n) \end{cases}$$

In particular, the above stated aggregation can assume a form of a standard averaging of the two actions, namely

$$y = \begin{cases} T_1 & \text{if } x \in \text{core}(A_1) \\ (T_1, T_2)/2 & \text{if } x \in \text{Shad}_+(A_1) \cup \text{Shad}_-(A_2) \\ & \vdots \\ T_n & \text{if } x \in \text{core}(A_n) \end{cases}$$

When we admit the induced fuzzy sets in place of the shadowed sets, the implicit expression of the uncertainty factor can be realized by a weighted averaging of the membership functions of the adjacent fuzzy sets as described below:

$$y = \begin{cases} T_1 & \text{if } x \in \text{core}(A_1) \\ \mu_1 T_1 + (1-\mu_2) T_2 & \text{if } x \in \text{Shad}_+(A_1) \cup \text{Shad}_-(A_2) \\ & \vdots \\ T_n & \text{if } x \in \text{core}(A_n) \end{cases}$$

Here μ_i is a membership value describing an 'activation' level of A_i.

One should reiterate the role of the confidence level(s) β used in this setting. They are associated with the selection of the relevant algorithm (method) T_i. More precisely,

$$P(\text{relevant } T_i \text{ selected}) = 1 - \beta$$

The choice of β can be thus completed as a result of an in-depth understanding and weighting of the consequences of the usage of T_i versus taking some other incorrect action.

The discussed method easily extends to the multidimensional case. Figure 9 shows its two-dimensional version involving two spaces (coordinates) $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ being partitioned with the use of two sets defined for each coordinate.

8. A Generalized Multiplexer

In this section, we elaborate on an interesting usage of the shadowed sets and fuzzy sets in the construction of a generalized multiplexer. This construct arises as a certain generalization of the well-known processing module being commonly used in the digital system design. To start with, it is instructive to discuss the essence of the system by considering the simplest possible scenario. We are interested in the two-variable mapping, denoted by $F(x, z)$ where the first variable of the mapping (x) is continuous whereas the second one (z) becomes granulated with the use of some sets. Without any loss in generality of the entire investigations, we may assume that $F(x,z)$ can read as follows:

$$y = F(x,z) = \begin{cases} \phi_1(x) \text{ if } z \in A_1 \\ \phi_2(x) \text{ if } z \in A_2 \end{cases}$$

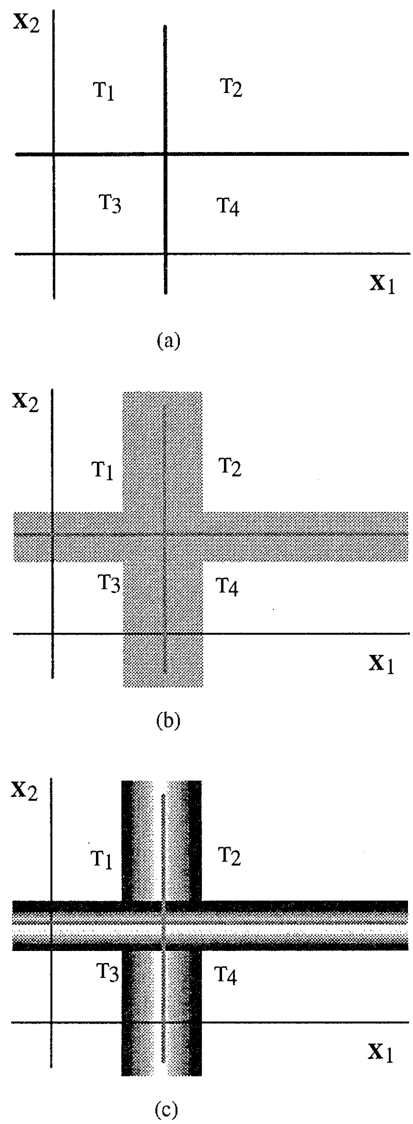

Fig. 9. Partition of the Cartesian product $\boldsymbol{X}_1 \times \boldsymbol{X}_2$ with the use of sets (a), shadowed sets (b), and fuzzy sets (c).

In a nutshell, the mapping can be realized with the aid of a multiplexer. With the only two sets used to granulate $\boldsymbol{Z}$, one can identify them with a single Boolean (two-valued) variable that completes switching between the two inputs, cf. Fig. 10.

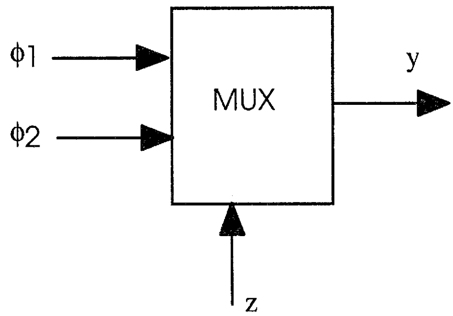

Fig. 10. A multiplexer structure of the two-variable mapping F.

The multiplexer architecture instantaneously generalizes to the cases involving more switching variables.

In the detailed example, we consider the relationship $y = F(x,z)$ with its components ϕ_1 and ϕ_2 defined as

$$\phi_1(x) = (x-1)^2 \text{ if } x \in [0,2]$$

$$\phi_2(x) = 3 - x \text{ if } x \in [2,4]$$

Moreover, $\boldsymbol{Z} = [-10, 10]$. The width of the interval of possible actions resulting from T_1 and T_2 is determined in the form

$$[\phi_1, \phi_2] = \max_{x \in X} \left| \phi_1(x), \phi_2(x) \right|$$

The performance index Q used to evaluate the behavior of various granulation mechanisms and the size of the information granules is a normalized standard sum of squared errors between $F(x,z)$ and $F(x,z')$ with z' being a noise-affected input. The data of the input variables are distributed uniformly over the two input spaces.

First, it is interesting to quantify how the performance index Q depends on the size of the shadow of the shadowed set and subsequently how this performance contrasts with the set-based multiplexer and another version designed based on the induced fuzzy set based counterpart. Figures 10(a), (b) and (c) visualize the values of the performance index versus D for selected values of the standard deviation of noise ($\sigma = 1, 2,$ and 4). In spite of small oscillations in the reported values of Q, the general tendency is obvious and intuitively appealing:

- For higher values of noise, the size of the shadows starts to grow. There is an optimal range of D's leading to the low values of the performance index.

- In all cases either the shadowed sets or fuzzy sets approach produces lower values of the performance index than those for set-based granulation (this situation is shown for $D = 0$).
- Fuzzy sets provide better results than shadowed sets. This is, again, not surprising: in the case of shadowed sets, the regions of their shadows come with a simple implementation of the averaging mechanism that produces constant results over the entire shadow. In the case of fuzzy sets, we come up with a continuous switching between the two local components (that is ϕ_1 and ϕ_2). This produces a continuous variability over the transition region.
- Flat regions that correspond to some ranges of D coincide with the optimal sizes of the shadows for a broad range of the confidence values.

It is also of interest to look at the behavior of the different versions of the multiplexer in terms of error and its overall distribution. The details are included in Fig. 12. To gain a better insight into the error distribution, the pertinent histograms of the error are included in Fig. 13. The performance of the multiplexer is succinctly summarized in the form of the normalized performance index regarded as a sum of the squared differences between $F(x, z)$ and $F(x, z')$ and averaged over all experimental data:

(a)	(b)	(c)
3.6233	2.3328	2.2625

The set-based granulation leads to the weakest performance while both shadowed and fuzzy set counterparts perform better; in the two latter cases the differences are less visible.

9. Conclusions

The paper has concentrated on some selected aspects of information granulation, its various ways of realization, and the ensuing properties resulting therein. It has been shown that the level of information granulation (namely, the size of information granules) and its further usage in modular algorithm is directly linked with the quality of input data to be used. The quantitative aspects of this dependency has been determined; we have found that an optimal level of information granulation is in a direct relationship with the intensity of random noise affecting the data. The paper includes a number of detailed relationships exemplifying these dependencies for some selected probabilistic characteristics of noisy data. By focusing on robustness (absorption) features of set-based information granulation, we have shown that this aspect can be easily described and further quantified in terms of the induced shadowed sets. The shadows of these constructs become instrumental in the delivery of some explicit expressions dealing with the uncertainty regions. Interestingly, one was able to derive direct dependencies between the size of the shadows of the shadowed sets

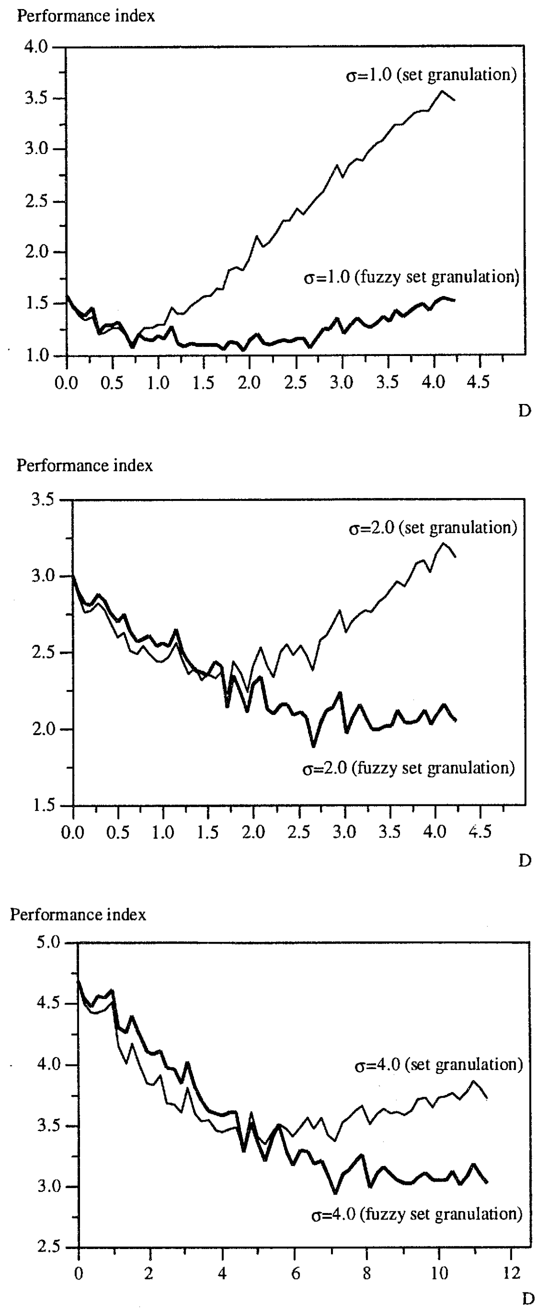

Fig. 11. The values of the performance index Q versus D for several selected levels of noise of the input data.

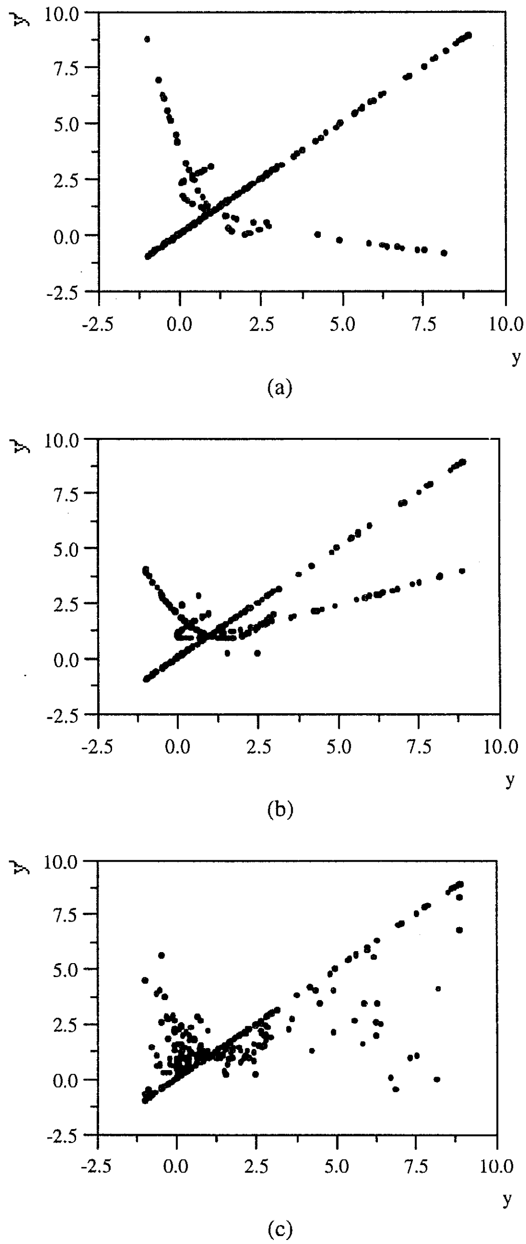

Fig. 12. Noise-free outputs of the mapping (y) versus results of the multiplexer (y'); the results concern 300 data points for Gaussian noise with $\sigma = 2$ and the confidence level of the confidence intervals equal to 0.15: (a) set based granulation, (b) shadowed set-oriented granulation, and (c) fuzzy set-oriented granulation.

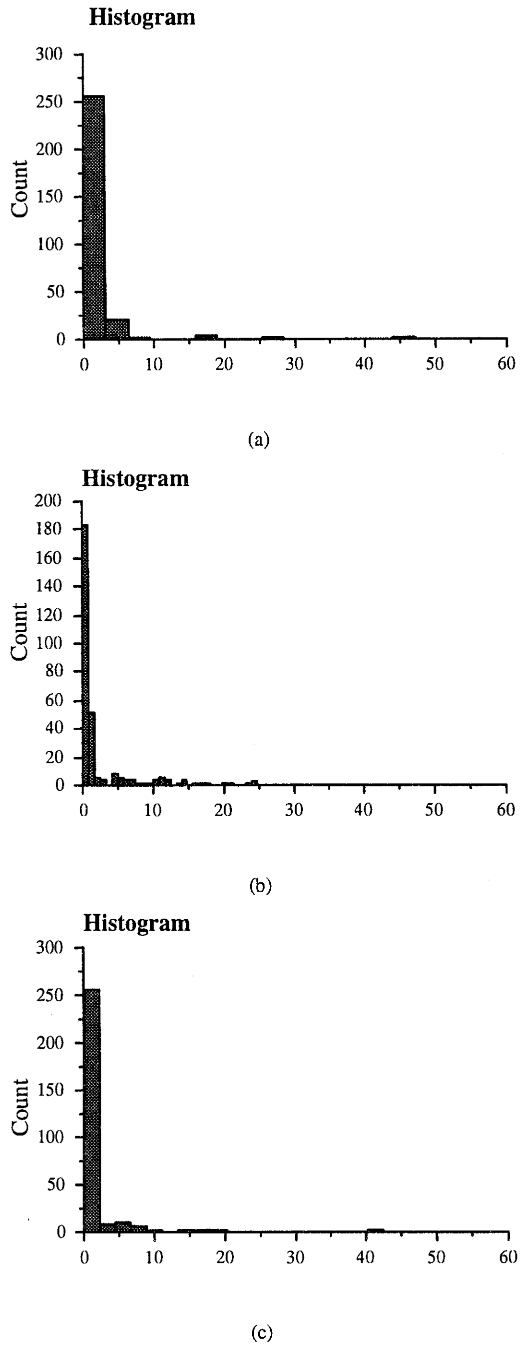

Fig. 13. Histograms of the error for the discussed versions of the multiplexers: (a) set based granulation, (b) shadowed set-oriented granulation, and (c) fuzzy set-oriented granulation.

and the variance of the noise affecting the input variable(s). We have also proposed another style of modeling that involves continuous as well as granulated variables. This bridges the classic ideas of continuous variable models and rule-based models that are predominant in the area of fuzzy models and other rule-oriented architectures. The concept of the generalized multiplexers gives another insight into the functioning of such models. Furthermore, we have investigated the use of shadowed sets as a crucial means of generating awareness about eventual relational behavior of the model in some regions of the data space. This mechanism is also essential to the better understanding of the data and the performance of the model. The relational effect has not been extensively studied yet, but its emergence becomes profound when we start analyzing experimental data that are not necessarily generated by well-known physical laws but rather originate from economical, social, and human-centered environments. As being a visibly focal point of various data mining pursuits, the role of relational and relational-functional modeling in such environments will be growing in importance in the near future. For the sake of completeness, one may indicate that the similar multimodel effect studied in (Pedrycz, 1996) exhibits a direct association with the relational phenomenon discussed here.

Acknowledgement

Support from the Natural Sciences and Engineering Research Council of Canada (NSERC) is gratefully acknowledged.

References

Kandel A. (1986): *Fuzzy Mathematical Techniques with Applications.* — Reading, MA: Addison-Wesley.

Lin T.Y. and Wildberger A.M. (Eds.) (1994): Proc. of 3-rd Int. Workshop *Rough Sets and Soft Computing RSSC'94*, San Jose State University, San Jose, CA.

Lin T.Y. and Cercone N. (Eds.) (1997): *Rough Sets and Data Mining. Analysis for Imprecise Data.* — Dordrecht: Kluwer.

Milanese M., Norton J., Piet-Lahanier H. and Walter E. (Eds.) (1996): *Bounding Approaches to System Identification.* — New York: Plenum Press.

Moore R.E. (1996): *Interval Analysis.* — Englewood Cliffs, NJ: Prentice Hall.

Moore R.E. (1979): *Methods of Interval Analysis.* — Philadelphia: SIAM.

Moore R.E. (Ed.) (1988): *Reliability in Computing: The Role of Interval Methods.* — New York: Academic Press.

Pawlak Z. (1982): *Rough sets.* — Int. J. Comp. Sci., Vol.11, No.4, pp.341–356.

Pawlak Z. (1991): *Rough Sets. Theoretical Aspects of Reasoning about Data.* — Dordrecht: Kluwer.

Pedrycz W. (1996): *Fuzzy multimodels.* — IEEE Trans. Fuzzy Syst., Vol.4, No.1, pp.139–148.

Pedrycz W. (1997): *Computational Intelligence: An Introduction.* — Boca Raton, Florida: CRC Press.

Pedrycz W. (1998): *Shadowed sets.* — IEEE Trans. Syst. Man Cybern., Part B, Vol.28, No.1, pp.103–109.

Pedrycz W. and Gomide F. (1998): *An Introduction to Fuzzy Sets: Analysis and Design.* — Cambridge, MA: MIT Press.

Sambuc R. (1975): *Fonctions Φ-flous. Application a l'aide au diagnostic en pathologie thyroidienne.* — Ph.D. Thesis, Marseille.

Skowron A. (1990): *The rough set theory and evidence theory.* — Fundamenta Informaticae, Vol.13, No.3, pp.245–262.

Slowinski R. (Ed.) (1992): *Intelligent Decision Support. Handbook of Applications and Advances of Rough Set Theory.* — Dordrecht: Kluwer.

Zadeh L.A. (1979): *Fuzzy sets and information granularity,* In: Gupta M.M., Ragade R.K., Yager R.R. (Eds.), Advances in Fuzzy Set Theory and Applications, Amsterdam: North Holland, pp.3–18.

Received: 9 February 1999
Revised: 7 April 1998