Journal of Uncertain Systems
Vol.1, No.1, pp.38-61, 2007
Online at: www.jus.org.uk

Granular Computing - The Emerging Paradigm

Witold Pedrycz

Department of Electrical & Computer Engineering
University of Alberta
Edmonton T6R 2G7 Canada
and
Systems Research Institute, Polish Academy of Sciences
Warsaw, Poland

Received 11 January 2007; Accepted 20 January 2007

Abstract

We provide an overview of Granular Computing - a rapidly growing area of information processing aimed at the construction of intelligent systems. We highlight the main features of Granular Computing, elaborate on the underlying formalisms of information granulation and discuss ways of their development. We also discuss the concept of granular modeling and present the issues of communication between formal frameworks of Granular Computing.

Keywords Granular Computing, fuzzy sets, rough sets, knowledge representation, granularity of information, granular communication.

1. Introduction

Fuzzy sets, interval analysis, rough sets are all constructs falling under the same umbrella of Granular Computing which has recently emerged as a coherent conceptual and algorithmic platform aimed at the representation and processing of information granules. The objective of this study is to provide a certain insight into the essence of Granular Computing being regarded as a coherent environment of manipulation of information granules, discuss the underlying formal ways of representing information granules, show pertinent design schemes and look at some application aspects of the area.

The paper is arranged into eight sections. In Section 2, we cover a concept of information granules and elaborate on their omnipresence in various endeavors of system modeling, control, decision–making, and classification. Next, identified and discussed are the key formalisms used to represent and process information granules such as sets, fuzzy sets, rough sets, and shadowed sets. Then we discuss ways of designing information granules; here to retain focus we concentrate on fuzzy sets and therefore concentrate on the determination of their membership functions. This selection is justified by the fact that there are numerous ways of membership function elicitation and such techniques could offer some general perspective at the formation of fuzzy sets. We address the issues of communication between various formal frameworks of Granular Computing. In particular, we note two essential dimensions that need to be taken into consideration, namely the level of granularity of the constructs and the underlying formalism within which such information granules are developed. A question of building granular models is also addressed. Concluding comments are offered in Section 8.

2. From information granules to Granular Computing

It becomes evident that information granules permeate human endeavors (Zadeh, 1996, 1997, 2005; Pedrycz, 2001; Bargiela and Pedrycz, 2003; Pedrycz and Gomide, 2007). No matter what problem is taken into account, we usually cast it into a certain conceptual framework of basic operational entities, which we regard to be of relevance to the problem formulation and problem solving. We formulate generic concepts adhering to some level of abstraction, carry out processing, and communicate the results to the external environment. Consider, for instance, image processing. In spite of the continuous progress in the area, a human being assumes a dominant and very much uncontested position when it comes to image understanding and image interpreting. Surely, we do not focus attention on individual pixels and process them as such but group them together into semantically meaningful constructs – familiar objects we deal with in everyday life. Such objects involve regions that consist of pixels or categories of pixels drawn together because of their proximity in the image, similar texture, color, etc. This remarkable and unchallenged ability of humans dwells on our effortless ability to construct information granules, manipulate them and arrive at sound conclusions. As another example, consider a collection of time series, say stock recordings. From our perspective we can describe them in a semi-qualitative manner by pointing at specific regions of such signals. Specialists can effortlessly interpret ECG signals. They distinguish some segments of such signals and interpret their combinations. Experts can interpret temporal readings of sensors and assess the status of the monitored system. Again, in all these situations, the individual samples of the signals are not the focal point of the analysis and the ensuing signal interpretation. We always granulate all phenomena (no matter if they are originally discrete or analog in their nature). Time is another important variable that is subjected to granulation. We use seconds, minutes, days, months, and years. Depending which specific problem we have in mind and who the user is, the size of information granules (time intervals) could vary quite dramatically. To the high level management time intervals of quarters of year or a few years could be meaningful temporal information granules on basis of which one develops any predictive model. For those in charge of everyday operation of a dispatching plant, minutes and hours could form a viable scale of time granulation. For the designer of high-speed integrated circuits and digital systems, the temporal information granules concern nanoseconds, microseconds, and perhaps microseconds. Even such commonly encountered and simple examples are convincing enough to lead us to ascertain that (a) information granules are the key components of knowledge representation and processing, (b) the level of granularity of information granules (their size, to be more descriptive) becomes crucial to the problem description and an overall strategy of problem solving, (c) there is no universal level of granularity of information; the size of granules is problem-oriented and user dependent.

What has been said so far touched a qualitative aspect of the problem. The challenge is to develop a computing framework within which all these representation and processing endeavors could be formally realized. The common platform emerging within this context comes under the name of Granular Computing. In essence, it is an emerging paradigm of information processing. While we have already noticed a number of important conceptual and computational constructs built in the domain of system modeling, machine learning, image processing, pattern recognition, and data compression in which various abstractions (and ensuing information granules) came into existence, Granular Computing becomes innovative and intellectually proactive in several different ways:

- It identifies the essential commonalities between the surprisingly diversified problems and technologies used there which could be cast into a unified framework we usually refer to as a granular world. This is a fully operational processing entity that interacts with the external world (that could be another granular or numeric world) by collecting necessary granular information and returning the outcomes of the granular computing
- With the emergence of the unified framework of granular processing, we get a better grasp as to the role of interaction between various formalisms and visualize a way in which they communicate.

- It brings together the existing formalisms of set theory (interval analysis), fuzzy sets, rough sets, etc. under the same roof by clearly visualizing that in spite of their visibly distinct underpinnings (and ensuing processing), they exhibit some fundamental commonalities. In this sense, Granular Computing establishes a stimulating environment of synergy between the individual approaches.
- By building upon the commonalities of the existing formal approaches, Granular Computing helps form heterogeneous and multifaceted models of processing of information granules by clearly recognizing the orthogonal nature of some of the existing and well established frameworks (say, probability theory coming with its probability density functions and fuzzy sets with their membership functions)
- Granular Computing fully acknowledges a notion of variable granularity whose range could cover detailed numeric entities and very abstract and general information granules. It looks at the aspects of compatibility of such information granules and ensuing communication mechanisms of the granular worlds.
- Interestingly, the inception of information granules is highly motivated. We do not form information granules without reason. Information granules are an evident realization of the fundamental paradigm of abstraction.

Granular Computing constitutes a unified conceptual and computing platform. Yet, it directly benefits from the already existing and well-established concepts of information granules formed in the setting of set theory, fuzzy sets, rough sets and others.

It is instructive to take a quick look at the fundamental technologies of information granulation and contrast their key features.

3. Fundamental formalisms of Granular Computing

There are several key formal frameworks contributing to Granular Computing and forming its algorithmic backbone. In what follows, we present their essential features.

3.1. Set theory and interval analysis

Sets are fundamental concepts of mathematics and science. Referring to the classic notes, set is described as "any multiplicity which can be thought of as one .. any totality of definite elements which can be bound up into a whole by means of a law" or being more descriptive "..any collection into a whole *M* of definite and separate objects m of our intuition or our thought" (Cantor, 1883, 1895). Likewise interval analysis ultimately dwells upon a concept of sets which in this case are collections of elements in the line of reals, say [a,b], [c,d],…etc. Multidimensional constructs are built upon Cartesian products of numeric intervals and give rise to computing with hyperboxes. Going back to the history, computing with intervals is intimately linked with the world of digital technology. One of the first papers in this area was published in 1956 by Warmus. Some other early research was done by Sunaga and Moore. This was followed by a wave of research in so-called interval mathematics or interval calculus. Conceptually, sets (intervals) are rooted in a two-valued logic with their fundamental predicate of membership ($\in$). Here holds an important isomorphism between the structure of two-valued logic endowed with its truth values (false-true) and set theory with sets being fully described by their characteristic functions. The interval analysis is a cornerstone of reliable computing which in turn is ultimately associated with digital computing in which any variable is associated with a finite accuracy (implied by the fixed number of bits used to represent numbers). This limited accuracy gives rise to a certain pattern of propagation of error of computing. For instance, addition of two intervals [a, b] and [c, d] leads to a broader interval in the form [a+c, b+d] (Hansen, 1975; Jaulin et al., 2001; Moore, 1966). Here the accumulation of uncertainty (or equivalently the decreased granularity of the result) depends upon the specific algebraic operation completed for given intervals. Table 2 summarizes four algebraic operations realized on numeric intervals A = [a, b] and B = [c, d].

algebraic operation	result
addition	[a+c, b+d]
subtraction	[a-d, b-c]
multiplication	[min(ac,ad, bc, bd), max(ac, ad, bc, bd)]
division	$[\min(\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d}), \max(\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d})]$ assumption: the interval [c, d] does not contain 0

Table 1. Arithmetic operations on numeric intervals A and B

Interestingly, intervals distributed uniformly in a certain space are at the center of any mechanism of analog-to-digital conversion; the higher the number of bits, the finer the intervals and the higher their number. The well known fundamental relationship states that with “n” bits we can build a collection of 2^n intervals of width $(b-a)/2^n$ for the original range of numeric values in [a, b]. Intervals offer a straightforward mechanism of abstraction: all elements lying within a certain interval become indistinguishable and therefore are treated as identical. In addition to algebraic manipulation, the area of interval mathematics embraces a wealth of far more advanced and practically relevant processing including differentiation, integral calculus as well as interval-valued optimization.

3.2. The role of fuzzy sets: a perspective of information granules

Fuzzy sets offer an important and unique feature of describing information granules whose contributing elements may belong with varying degrees of membership (belongingness). This helps us describe the concepts that are commonly encountered in real word. The notions such as *low* income, *high* inflation, *small* approximation error and many others are examples of concepts to which the yes-no quantification does not apply or becomes quite artificial and restrictive. We are cognizant that there is no way of quantifying the Boolean boundaries as there are a lot of elements whose membership to the concept is only partial and quite different from 0 and 1.

The binary view of the world supported by set theory and two-valued logic has been vigorously challenged by philosophy and logic. The revolutionary step in logic was made by Lukasiewicz with his introduction of three and afterwards multivalued logic. It took however more decades to dwell on the ideas of the non-Aristotelian view of the world before fuzzy sets were introduced. This happened in 1965 with the publication of the seminal paper on fuzzy sets by Zadeh (1965). Refer also to other influential papers by Zadeh (1996, 1997, 1999, 2005). The concept of fuzzy set is surprisingly simple and elegant. Fuzzy set A captures its elements by assigning them to it with some varying degrees of membership. A so called membership function is a vehicle that quantifies different degrees of membership. The higher the degree of membership A(x), the stronger is the level of belongingness of this element to A (Gottwald, 2005; Zimmermann, 1996; Pedrycz and Gomide, 1998; Pedrycz and Gomide, 2007).

The obvious yet striking difference between sets (intervals) and fuzzy sets lies in the notion of partial membership supported by fuzzy sets. In fuzzy sets, we discriminate between elements that are “typical” to the concept and those of borderline character. Information granules such as *high* speed, *warm* weather, *fast* car are examples of information granules falling under this category can be conveniently represented by fuzzy sets. As we cannot specify a single, well-defined element that forms a solid border between full belongingness and full exclusion, fuzzy sets offer an appealing alternative and a practical solution to this problem. Fuzzy sets with their smooth transition boundaries form an ideal vehicle to capture the notion of partial membership. In this sense information granules formalized in the language of fuzzy sets support a vast array of human-centric pursuits. They are predisposed to play a vital role when interfacing human to intelligent systems.

In problem formulation and problem solving, fuzzy sets emerge in two fundamentally different ways,

explicit. Here, they typically pertain to some generic and fairly basic concepts we use in our communication and description of reality. There is a vast amount of examples as such concepts being commonly used every day, say *short* waiting time, *large* dataset, *low* inflation, *high* speed, *long* delay, etc. All of them are quite simple as we can easily capture their meaning. We can easily identify a universe of discourse over which such variable are defined. For instance, this could be time, number of records, velocity, and alike.

implicit Here we are concerned with more complex and inherently multifaceted concepts and notions where fuzzy sets could be incorporated into the formal description and quantification of such problems yet not in so instantaneous manner. Some examples could include concepts such as "*preferred* car", "*stability* of the control system", "*high performance* computing architecture", "*good convergence* of the learning scheme", *strong* economy, etc. All of these notions incorporate some components that could be quantified with the use of fuzzy sets yet this translation is not that completely straightforward and immediate as it happens for the category of the explicit usage of fuzzy sets. For instance, the concept of "*preferred* car" is evidently multifaceted and may involve a number of essential descriptors that when put together are really reflective of the notion we have in mind. For instance, we may involve a number of qualities such as speed, economy, reliability, depreciation, maintainability, and alike. Interestingly, each of these features could be easily rephrased in simpler terms and through this process at some level of this refinement phase we may arrive at fuzzy sets that start to manifest themselves in an explicit manner.

As we stressed, the omnipresence of fuzzy sets is surprising. Even going over any textbook or research monograph, not mentioning newspapers and magazines, we encounter a great deal of fuzzy sets coming in their implicit or explicit format. Table 2 offers a handful of selected examples

p. 65: *small* random errors in the measurement vector…
p. 70: The success of the method depends on whether the first initial guess is already *close enough* to the global minimum…
p. 72: Hence, the convergence region of a numerical optimizer will be *large*

F. van der Heijden et al., *Classification, Parameter Estimation and State Estimation*, J. Wiley, 2004, Chichester.

p. 50: validation costs are *high* for *critical systems*
p. 660: …A *high* value for fan-in means that X is *highly coupled* to the rest of the design and changes to X will have extensive knock-on effect. A *high* value for fan-out suggests that the overall complexity of X may be *high* because of the complexity of control logic needed to coordinate the called components.
… Generally, the *larger* the size of the code of a component, the more *complex* and error-prone the component is likely to be…
… The *higher* the value of the Fog index, the more difficult the document is to understand

I. Sommerville, *Software Engineering*, 8th edition, Addison-Wesley, 2007, Harlow.

Table 2. Examples of concepts whose description and processing invokes the use of the technology of fuzzy sets and Granular Computing

From the optimization standpoint, the properties of continuity and commonly encountered differentiability of the membership functions becomes a genuine asset. We may easily envision situations where those information granules incorporated as a part of the neurofuzzy system are subject to optimization – hence the differentiability of their membership functions becomes of critical relevance. What becomes equally important is the fact that fuzzy sets bridge numeric and symbolic concepts. On one hand, fuzzy set can be treated as some symbol. We can regard it as a single conceptual entity by assigning to it some symbol, say L (for *low*). In the sequel, it could be processed as a purely symbolic entity. On the other hand, a fuzzy set comes with a numeric membership function and these membership grades could be processed in a numeric fashion.

Fuzzy sets can be viewed from several fundamentally different standpoints. Here we emphasize the three of them that play a fundamental role in processing and knowledge representation.

as a enabling processing technology of some universal character and of profound human-centric character

Fuzzy sets build upon the existing information technologies by forming a user-centric interface using which one could communicate essential design knowledge thus guiding problem solving and making it more efficient. For instance, in signal processing and image processing we might incorporate a collection of rules capturing specific design knowledge about filter development in a certain area. Say, "if the level of noise is *high*, consider using a *large* window of averaging" In control engineering, we may incorporate some domain knowledge about the specific control objectives. For instance, "if the constraint of fuel consumption is *very important*, consider settings of a PID controller producing *low* overshoot" Some other examples of highly representative human-centric systems concern those involving (a) construction and usage of relevance feedback in retrieval, organization and summarization of video and images, (b) queries formulated in natural languages, (c) summarization of results coming as an outcome some query.

Secondly, there are unique areas of applications in which fuzzy sets form a methodological backbone and deliver the required algorithmic setting. This concerns fuzzy modeling in which we start with collections of information granules (typically realized as fuzzy sets) and construct a model as a web of links (associations) between them. This approach is radically different from the numeric, function-based models encountered in "standard" system modeling. Fuzzy modeling emphasizes an augmented agenda in comparison with the one stressed in numeric models. While we are still concerned with the accuracy of the resulting model, its interpretability and transparency becomes of equal and sometimes even higher relevance.

as an efficient computing framework of global character

Rather than processing individual elements, say a single numeric datum, an encapsulation of a significant number of the individual elements that is realized in the form of some fuzzy sets, offers immediate benefits of joint and orchestrated processing. Instead of looking at the individual number, we embrace a more general point of view and process a entire collection of elements represented now in the form of a single fuzzy set. This effect of a *collective* handling of individual elements is seen very profoundly in so-called fuzzy arithmetic. The basic constructs here are fuzzy numbers. In contrast to single numeric quantities (real numbers) fuzzy numbers represent collections of numbers where each of them belongs to the concept (fuzzy number) to some degree. These constructs are then subject to processing, say addition, subtraction, multiplication, division, etc. Noticeable is the fact that by processing fuzzy numbers we are in fact handling a significant number of individual elements at the same time. Fuzzy numbers and fuzzy arithmetic provide an interesting advantage over interval arithmetic (viz. arithmetic in which we are concerned with intervals – sets of numeric values). Intervals come with abrupt boundaries as elements can belong or are excluded from the given set. This means, for example, that any gradient-based techniques of optimization invoked when computing solutions become very limited: the derivative is equal to zero with an exception at the point where the abrupt boundary is located.

fuzzy sets as a vehicle of raising and quantifying awareness about granularity of outcomes

Fuzzy sets form the results of granular computing. As such they convey a global view at the elements of the universe of discourse over which they are constructed. When visualized, the values of the membership function describe a suitability of the individual points as compatible (preferred) with the solution. In this sense, fuzzy sets serve as a useful visualization vehicle: when displayed, the user could gain an overall view at the character of solution (regarded as a fuzzy set) and make a final choice. Note that this is very much in line with the idea of the human-centricity: we present the user with all possible results however do not put any pressure as to the commitment of selecting a certain numeric solution.

fuzzy sets as a mechanism realizing a principle of the least commitment

As the computing realized in the setting of granular computing returns a fuzzy set as its result, it could be effectively used to realize a principle of the least commitment. The crux of this principle is to use fuzzy set as a mechanism of making us cognizant of the quality of obtained result. Consider a fuzzy set being a result of computing in some problem of multiphase decision-making. The fuzzy set is defined over various alternatives and associates with them the corresponding degrees of preference, see Figure 1. If there are several alternatives with very similar degrees of membership, this serves as a clear indicator of uncertainty or hesitation as to the making a decision. In other words, in light of the form of the generated fuzzy set, we do not intend to commit ourselves to making any decision (selection of one of the alternatives) at this time. Our intent would be to postpone decision and collect more evidence. For instance, this could involve further collecting of data, soliciting expert opinion, and alike. Based on this evidence, we could continue with computing and evaluate the form of the resulting fuzzy set. It could well be that the collected evidence has resulted in more specific fuzzy set of decisions on basis of which we could either still postpone decision and keep collecting more evidence or proceed with decision-making. Thus the principle of the least commitment offers us an interesting and useful guideline as to the mechanism of decision-making versus evidence collection.

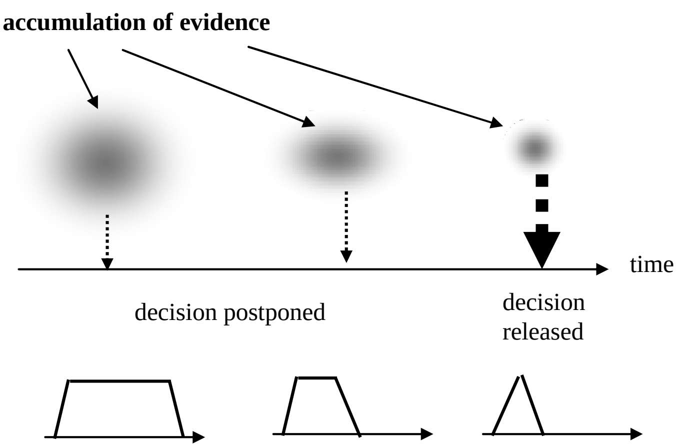
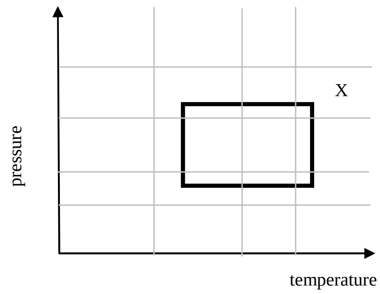


Figure 1. An essence of the principle of the least commitment; the decision is postponed until the phase where there is enough evidence accumulated and the granularity of the result becomes specific enough. Shown are also examples of fuzzy sets formed at successive phases of processing that become more specific along with the increased level of evidence.

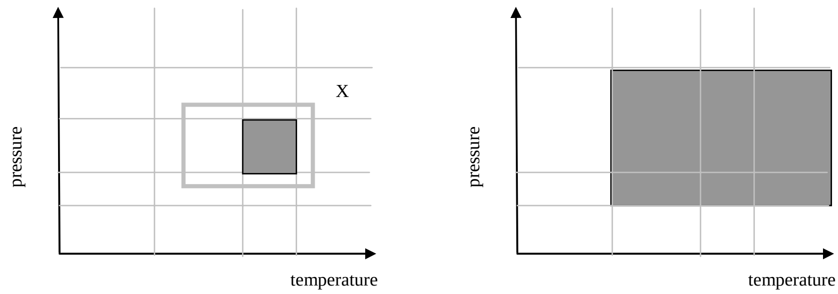
3.3. Rough sets

The description of information granules completed with the aid of some vocabulary is usually imprecise. Intuitively, such description may lead to some approximations, called lower and upper bounds. This is the essence of rough sets introduced by Pawlak (1982; 1991); refer also to Polkowski and Skowron (1998) and Skowron (1989). Interesting generalizations, conceptual insights, and algorithmic investigations are offered in a series of fundamental papers by Pawlak and Skowron (2007).

To explain the concept of rough sets and show what they are to offer in terms of representing information granules, we us use an illustrative example. Consider a description of environmental conditions expressed in terms of temperature and pressure. For each of these factors, we fix several ranges of possible values where each of such ranges comes with some interpretation such as “values below”, “values in-between”, “values above”, etc. By admitting such selected ranges in both variables, we construct a grid of concepts formed in the Cartesian product of the spaces of temperature and pressure, refer to Figure 2. In more descriptive terms, this grid forms a vocabulary of generic terms using which we would like to describe all new information granules.



(a)



(b)

Figure 2. A collection of vocabulary and their use in the problem description. Environmental conditions X result in some interval of possible values (a). In the sequel, this gives rise to the concept of a rough set with the roughness of the description being captured by the lower and upper bounds (approximations) as illustrated in (b).

Now let us consider that the environmental conditions monitored over some time have resulted in some values of temperature and pressure ranging in-between some lower and upper bound as illustrated in Figure 2. Denote this result by X. When describing it in the terms of the elements of the vocabulary, we end up with a collection of elements that are fully included in X. They form a lower bound of description of X when being completed in presence of the given vocabulary. Likewise, we may identify elements of the vocabulary that have a nonempty overlap with X and in this sense constitute an upper bound of the description of the given environmental conditions. Along with the vocabulary, the description forms a certain rough set.

It is interesting to note that the vocabulary used in the above construct could comprise information granules being expressed in terms of any other formalism, say fuzzy sets. Quite often we can encounter constructs like rough fuzzy sets and fuzzy rough sets in which both fuzzy sets and rough sets are put together (Dubois and Prade, 1990).

3.5. Shadowed sets

Fuzzy sets are associated with the collections of numeric membership grades. Shadowed sets (Pedrycz, 1998; 2005) are based upon fuzzy sets by forming a more general and highly synthetic view at the numeric concept of membership. Using shadowed sets, we quantify numeric membership values into three categories: complete belongingness, complete exclusion and unknown (which could be also conveniently referred to as don't know condition or a *shadow*). A graphic illustration of a shadowed set along with the principles of sets and fuzzy sets is schematically shown in Figure 3. This helps us contrast these three fundamental constructs of information granules.

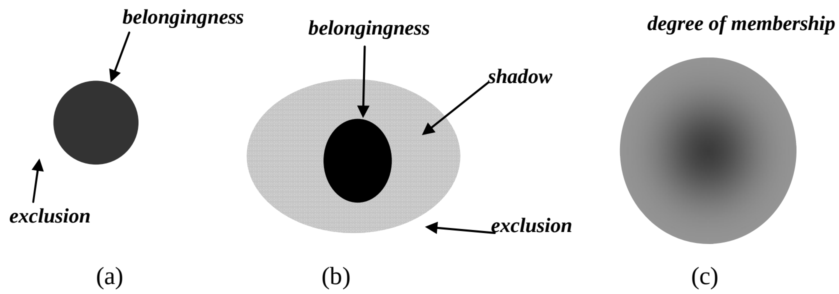


Figure 3. A schematic view at sets (a), (b) shadowed sets, and fuzzy sets (c).
Shadowed sets reveal interesting linkages between fuzzy sets and sets.

In a nutshell, shadowed sets can be regarded as a general and far more concise representation of a fuzzy set that could be of particular interest when dealing with further computing (in which case we could come up with substantial reduction of the overall processing effort).

Fuzzy sets help describe and quantify concepts of continuous boundaries. By introducing an α-*cut*, we can convert a fuzzy set into a set. By choosing a threshold level (α) that is high enough, we admit elements whose membership grades are meaningful (as viewed from the standpoint of the used threshold). The fact that an α -cut transforms a fuzzy set into a set, could create the impression that any fuzzy set can be made equivalent to some set. This point of view is highly deceptive. In essence, by building any α -cut, we elevate some membership grades to 1 (full membership) and eliminate other with lower membership grades (total exclusion). Surprisingly, this process does not take into account the distribution of elements with partial membership so this effect cannot be quantified in the resulting construct. The idea of shadowed sets aims at alleviating this problem by constructing regions of complete ignorance about membership grades. In essence, a shadowed set $A^{\sim}$ induced by a given fuzzy set A defined in **X** is an interval-valued set in **X** which maps elements of **X** into 0, 1, and the entire unit interval, that is [0,1]. Formally, $A^{\sim}$ is a mapping:

$$A^{\sim} : \mathbf{X} \rightarrow \{ 0, 1, [0,1]\} \quad (1)$$

Given $A^{\sim}(x)$, the two numeric values (0 and 1) take on a standard interpretation: 0 denotes complete exclusion from $A^{\sim}$, while 1 stands for complete inclusion in *A*. $A^{\sim}(x)$ equal to [0,1] represents a complete ignorance – nothing is known about the membership of *x* in $A^{\sim}$: we *neither* confirm its belongingness to $A^{\sim}$ *nor* commit to its exclusion. In this sense, such as "x" is the most "questionable" point and should be treated as such (e.g., this outcome could trigger some action to analyze this element in more detail, exclude it from further analysis, etc.). The name *shadowed set* is a descriptive reflection of a set that comes with "shadows" positioned around the edges of the characteristic function, as illustrated in Figure 4.

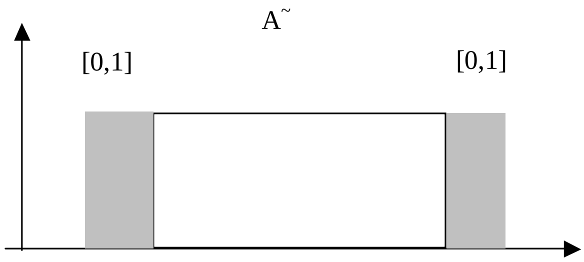


Figure 4. A shadowed set $A^{\sim}$. Note “shadows” produced at the edges of the characteristic function.

Shadowed sets are isomorphic with a three-valued logic. Operations on shadowed sets are the same as in this logic. The underlying principle is to retain the vagueness of the arguments (shadows of the shadowed sets being used in the aggregation). The following tables capture the description of the operators on shadowed sets:

Union

0	0	1	[0,1]
1	1	1	1
[0,1]	[0,1]	1	[0,1]
	0	1	[0,1]

Intersection

0	0	0	0
1	0	1	[0,1]
[0,1]	0	[0,1]	[0,1]
	0	1	[0,1]

Complement

0	1
1	0
[0,1]	[0,1]

From the design point of view, shadowed sets are induced by fuzzy sets, and in this setting their role is to help interpret results given in the form of fuzzy sets and to reduce computational overhead. Since shadowed sets do not focus on detailed membership grades and process only 0, 1, and ½ (considering that the numeric value of ½ is used to code the shadow), all processing is very simple and computationally appealing.

Given the underlying motivation, the development of shadowed sets starts from a given fuzzy set. The underlying criterion governing this transformation is straightforward: maintain a balance of uncertainty in the sense that, while reducing low membership grades to zero and bringing high membership grades to 1, maintain the overall balance of change in membership grades. The changes of membership grades to 0 and 1 are compensated for by the construction of the shadow that “absorbs” the previous elimination of partial membership at low and high ranges of membership. This design principle for a unimodal fuzzy set is illustrated in Figure 5. The transformation is guided by the value of threshold β; more specifically, we are concerned with two individual thresholds, namely, β and $1-\beta$.

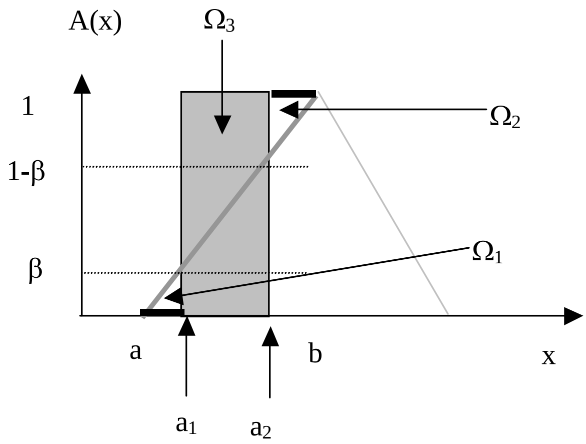


Figure 5. Induced shadowed set. The elimination of regions of partial membership is counterbalanced by the formation of shadows "absorbing" the reduction realized in the region of partial membership grades.

The retention of balance translates into the following dependency:

$$\Omega_1 + \Omega_2 = \Omega_3 \tag{2}$$

where the corresponding regions are illustrated in Figure 5. Note that we are dealing with the increasing and decreasing portions of the membership functions separately. The integral form of the above relationship is given as

$$\int_{a}^{a_1} A(x)dx + \int_{a_2}^{b} (1 - A(x))dx = \int_{a_1}^{a_2} dx \tag{3}$$

For its detailed interpretation, refer again to Figure 5. A certain threshold value of β, $\beta \in [0,1/2)$ that satisfies this expression is treated as a solution to the problem. Based on this result, we form a shadow of the shadowed set. In the case of commonly encountered membership functions, the optimal value of β can be determined in an analytical manner. For the triangular membership function, we consider each segment (the increasing and decreasing portion of the membership function) separately and focus on the linearly increasing portion of the membership function governed by an expression of the form $(x-a)/(b-a)$. Simple calculations reveal that the cutoff points a_1 and a_2 are equal to $a+\beta$ (b-a) and a $+(1-\beta)$(b-a). Subsequently, the resulting optimal value of β is equal to $\frac{2^{3/2}-2}{2} = 0.4142$. Similarly, when dealing with a nonlinear membership function such that $A(x) = \sqrt{\frac{x-a}{b-a}}$ in $x \in [a,b]$ and zero outside this interval we get $a_1 = a + \beta^2(b-a)$ and $a_2 = a + (1-\beta)^2(b-a)$. The only root that satisfies the requirements imposed on the threshold values is equal to 0.405.

4. The design of information granules

In this chapter, we focus on the development of fuzzy sets by presenting various ways of forming fuzzy sets and determining their membership functions. The subject of elicitation and interpretation of fuzzy sets (membership functions) is of paramount relevance from the conceptual, algorithmic, and application-oriented standpoints. There is a significant diversity of the methods that support the construction of membership functions. In general, one can clearly distinguish between user-driven and data driven approaches with a number of techniques that share some features specific to both data- and user-driven techniques and hence are located somewhere in-between. The determination of membership functions has been a debatable issue for a long time almost since the very inception of fuzzy sets. In contrast to interval analysis and set theory where the estimation of bounds of the interval constructs has not attracted a great deal of attention and seemed to be taken for granted, an estimation of membership degrees became essential and over time has led us to sound, suite well justified and algorithmically appealing estimation techniques.

4.1. Semantics of information granules: some general insights

Fuzzy sets are constructs that come with a well defined meaning. They capture the semantics of the framework they intend to operate within. Fuzzy sets are the building conceptual blocks (generic constructs) that are used in problem description, modeling, control, and pattern classification tasks. Before discussing specific techniques of membership function estimation, it is worth casting the overall presentation in a certain context by emphasizing the aspect of the use of a finite number of fuzzy sets leading to some essential vocabulary reflective of the underlying domain knowledge. In particular, we are concerned with the related semantics, calibration capabilities of membership functions and the locality of fuzzy sets.

The limited capacity of a short term memory, as identified by Miller suggests that we could easily and comfortably handle and process 7 ± 2 items. This implies that the number of fuzzy sets to be considered as meaningful conceptual entities should be kept at the same level. The observation sounds reasonable — quite commonly in practice we witness situations in which this holds. For instance, when describing linguistically quantified variables, say error or change of error, we may use seven generic concepts (descriptors) labeling them as positive *large*, positive *medium*, positive *small*, *around* zero, negative *small*, negative *medium*, negative *large*. When characterizing speed, we may talk about its quite intuitive descriptors such as *low*, *medium* and *high* speed. In the description of an approximation error, we may typically use the concept of a *small* error around a point of linearization (in all these examples, the terms are indicated in italics to emphasize the granular character of the constructs and the role being played there by fuzzy sets). While embracing very different tasks, all these descriptors exhibit a striking similarity. All of them are information granules, not numbers (whose descriptive power is very much limited). In modular software development when dealing with a collection of modules (procedures, functions and alike), the list of their parameters is always limited to a few items which is again a reflection of the limited capacity of the short term memory. The excessively long parameter list is strongly discouraged due to the possible programming errors and rapidly increasing difficulties of an effective comprehension of the software structure and ensuing flow of control.

In general, the use of an excessive number of terms does not offer any advantage. To the contrary: it remarkably clutters our description of the phenomenon and hampers further effective usage of such concepts we intend to establish to capture the essence of the domain knowledge. With the increase in the number of fuzzy sets, their semantics becomes also negatively impacted. Fuzzy sets may be built into a hierarchy of terms (descriptors) but at each level of this hierarchy (when moving down towards higher specificity that is an increasing level of detail), the number of fuzzy sets is kept at a certain limited level.

While fuzzy sets capture the semantics of the concepts, they may require some calibration depending upon the specification of the problem at hand. This flexibility of fuzzy sets should not be treated as any shortcoming but rather viewed as a certain and fully exploited advantage. For instance, a term *low* temperature comes with a clear meaning yet it requires a certain calibration depending upon the environment and the context it was put into. The concept of *low* temperature is used in different climate zones and is of relevance in any communication between people yet for each of the community the meaning of the term is different thereby requiring some calibration. This could be realized e.g., by shifting the membership function along the universe of discourse of temperature, affecting the universe of discourse by some translation, dilation and alike. As a communication means, linguistic terms are fully legitimate and as such they appear in different settings. They require some refinement so that their meaning is fully understood and shared by the community of the users.

When discussing the methods aimed at the determination of membership functions or membership grades, it is worthwhile to underline the existence of the two main categories of approaches being reflective of the origin of the numeric values of membership. The first one is reflective of the domain knowledge and opinions of experts. In the second one, we consider experimental data whose global characteristics become reflected in the form and parameters of the membership functions. In the first group we can refer to the pairwise comparison (known also as a Saaty's approach) as one of the representative examples while fuzzy clustering is usually presented as a typical example of the data-driven method of membership function estimation. In what follows, we elaborate on several representative methods which will help us appreciate the level and flexibility of fuzzy sets.

5. The development of fuzzy sets

Fuzzy sets come with a number of algorithmic ways of forming their membership functions. In this section, we elaborate on a number of representative methods of their estimation.

5.1. Fuzzy set as a descriptor of feasible solutions

The aim of the method is to relate membership function to the level of feasibility of individual elements of a family of solutions associated with the problem at hand. Let us consider a certain function f(x) defined in $\boldsymbol{\Omega}$, that is f: $\boldsymbol{\Omega} \rightarrow \mathbf{R}$. where $\boldsymbol{\Omega} \subset \mathbf{R}$. Our intent is to determine its maximum, namely $x^{opt} = \arg\max_x f(x)$. On a basis of the values of f(x), we can form a fuzzy set A describing a collection of feasible solutions that could be labeled as optimal. Being more specific, we use the fuzzy set to represent an extent (degree) to which some specific values of "x" could be sought as potential (optimal) solutions to the problem. Taking this into consideration, we relate the membership function of A with the corresponding value of f(x) cast in the context of the boundary values assumed by "f". For instance, the membership function of A could be expressed in the following form

$$A(x) = \frac{f(x) - f_{min}}{f_{max} - f_{min}} \tag{4}$$

The boundary conditions are straightforward: $f_{min} = \min_x f(x)$ and $f_{max} = \max_x f(x)$ where the minimum and the maximum are computed over $\boldsymbol{\Omega}$. For other values of "x" where f attains is maximal value, A(x) is equal 1 and around this point, the membership values are reduced when "x" is likely to be a solution to the problem $f(x) < f_{max}$. The form of the membership function depends upon the character of the function under consideration. The following examples illustrate the essence of the construction of membership functions.

5.2. Fuzzy set as a descriptor of the notion of typicality

Fuzzy sets address an issue of gradual *typicality* of elements to a given concept. They stress the fact that there are elements that fully satisfy the concept (are typical for it) and there are various elements that are allowed only with partial membership degrees. The form of the membership function is reflective of the semantics of the concept. Its details could be captured by adjusting the parameters of the membership function or choosing its form depending upon experimental data. For instance, consider a fuzzy set of squares. Formally, a rectangle includes a square shape as its special example when the sides are equal, a = b, Figure 6. What if a = b +ε where ε is a very small positive number? Could this figure be sought as a square? It is very likely that we could agree with this notion. Perhaps the membership value of the corresponding membership function could be equal to 0.99. Our perception, which comes with some level of tolerance to imprecision, does not allow us to tell apart this figure from the ideal square, Figure 6.

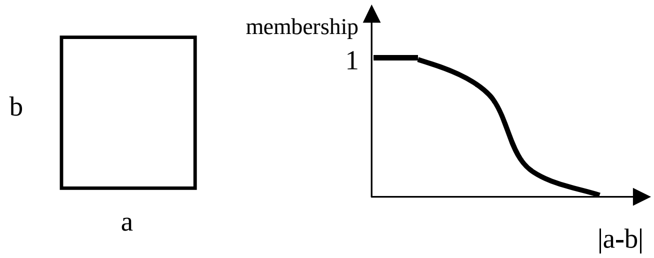


Figure 6. Perception of geometry of squares and its quantification in the form of membership function of the concept of fuzzy square.

Higher differences between "a" and "b" could result in lower values of the membership function. The definition of the fuzzy set square could be formed in a number of ways. Prior to the definition or even visualization of the membership function, it is important to formulate a space over which it will be defined. There are several intuitive alternatives worth considering:

(a) for each pair of values of the sides (a and b), collect an experimental assessment of membership of the rectangle to the category of squares. Here the membership function is defined over a Cartesian space of the spaces of lengths of sides of the rectangle. While selecting a form of the membership we require that it assumes values at a = b and is gradually reduced when the arguments start getting more different.

(b) we can define an absolute distance between "a" and "b" , |a-b| and form a fuzzy set over this space **X**; **X** = {x| x = |a-b|} $\mathbf{X} \subset \mathbf{R}_+$. The semantic constraints translate into the condition of A(0) = 1. For higher values of "x" we may consider monotonically decreasing values of A.

(c) we can envision ratios of a and b x = a/b and construct a fuzzy set over the space of $\mathbf{R}_+$ such that **X** = {x| x = a/b}. Obviously, we require here that A(1) =1. We also anticipate lower values of membership grades when moving to the left and to the right from x=1. Note that the membership function could be asymmetric so we allow for different membership values for the same length of the sides, say a =6, b= 5 and a =6 and b =5 (the effect could be quite apparent due to the visual effects when perceiving geometric phenomena). The previous model of **X** as outlined in (a) cannot capture this effect.

Once the form of the membership function has been defined, it could be further adjusted by modifying the values of its parameters on a basis of some experimental findings. They come in the form of ordered triples or pairs, say (a, b, μ), (a/b, μ) or (|a-b|, μ) depending on the previously accepted definition of the universe of discourse. The membership values μ are those available from the expert offering an assessment of the likeness of the corresponding geometric figure.

5.3. Vertical and horizontal schemes of membership estimation

The vertical and horizontal modes of membership estimation are two standard approaches used in the determination of fuzzy sets. They reflect distinct ways of looking at fuzzy sets whose membership functions at some finite number of points are quantified by experts. In the horizontal approach we identify a collection of elements in the universe of discourse **X** and request that an expert answers the question

-does x belong to concept A?

The answers are expected to come in a binary (yes-no) format. The concept A defined in **X** could be any linguistic notion, say *high* speed, *low* temperature, etc. Given "n" experts whose answers for a given point of **X** form a mix of yes-no replies, we count the number of "yes" answers and compute the ratio of the positive answers (p) versus the total number of replies(n), that is p/n. This ratio (likelihood) is treated as a membership degree of the concept at the given point of the universe of discourse. When all experts accept that the element belongs to the concept, then its membership degree is equal to 1. Higher disagreement between the experts (quite divided opinions) results in lower membership degrees. The concept A defined in **X** requires collecting results for some other elements of **X** and determining the corresponding ratios as outlined in Figure 7.

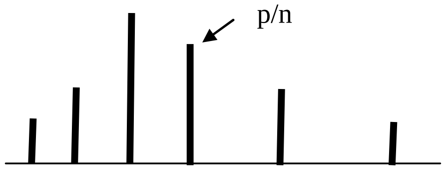


Figure 7. A horizontal method of the estimation of the membership function; observe a series of estimates determined for selected elements of **X**. Note also that the elements of **X** need not to be evenly distributed.

If replies follow some, e.g., binomial distribution, then we could y determine a confidence interval of the individual membership grade. The standard deviation of the estimate of the positive answers associated with the point x, denoted here by σ is given in the form

$$\sigma = \sqrt{\frac{p(1-p)}{n}} \tag{5}$$

The associated confidence interval which describes a range of membership values is then determined as

$$[p- \sigma, p+ \sigma] \tag{6}$$

In essence, when the confidence intervals are taken into consideration, the membership estimates become intervals of possible membership values and this leads to the concept of so-called interval-valued fuzzy sets. By assessing the width of the estimates, we could control the execution of the experiment: when the ranges are too long, one could re-design the experiment and monitor closely the consistency of the responses collected in the experiment.

The vertical mode of membership estimation is concerned with the estimation of the membership function by focusing on the determination of the successive α -cuts. The experiment focuses on the unit interval of membership grades. The experts involved in the experiment are asked the questions of the form

-what are the elements of **X** which belong to fuzzy set A at degree not lower than α?

where α is a certain level (threshold) of membership grades in [0,1]. The essence of the method is illustrated in Figure 8. Note that the satisfaction of the inclusion constraint is obvious: we envision that for higher values of α, the expert is going to provide more limited subsets of **X**; the vertical approach leads to the fuzzy set by combining the estimates of the corresponding α -cuts. Given the nature of this method, we are referring to the collection of random sets as these estimates appear in the successive stages of the estimation process.

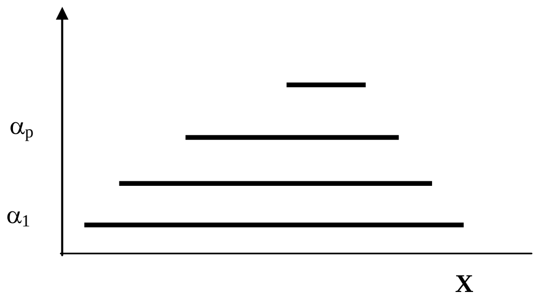


Figure 8. A vertical approach of membership estimation through the reconstruction of a fuzzy set through its estimated α -cuts.

The elements are identified by the expert as they form the corresponding α -cuts of A. By repeating the process for several selected values of α we end up with the α -cuts and using them we reconstruct the fuzzy set. The simplicity of the method is its genuine advantage. Like in the horizontal method of membership estimation, a possible lack of continuity is a certain disadvantage one has to be aware of. Here the selection of suitable levels of α needs to be carefully investigated. Similarly, an order at which different levels of α are used in the experiment could impact the estimate of the membership function.

5.4. Saaty's priority method of pairwise membership function estimation

The starting point of the estimation process are entries of the reciprocal matrix which are obtained through collecting results of pairwise evaluations offered by an expert, designer or user (depending on the character of the task at hand). Prior to making any assessment, the expert is provided with a finite scale with values spread in-between 1 to 7. Some other alternatives of the scales such as those involving 5 or 9 levels could be sought as well. If x_i is strongly preferred over x_j when being considered in the context of the fuzzy set whose membership function we would like to estimate, then this judgment is expressed by assigning high values of the available scale, say 6 or 7. If we still sense that x_i is preferred over x_j yet the strength of this preference is lower in comparison with the previous case, then this is quantified using some intermediate values of the scale, say 3 or 4. If no difference is sensed, the values close to 1 are the preferred choice, say 2 or 1. The value of 1 indicates that x_i and x_j are equally preferred. On the other hand, if x_j is preferred over x_i, the corresponding entry assumes values below one. Given the reciprocal character of the assessment, once the preference of x_i over x_j has been quantified, the inverse of this number is plugged into the entry of the matrix that is located at the (j,i)-th coordinate. As indicated earlier, the elements on the main diagonal are equal to 1. Next the maximal eigenvalue is computed along with its corresponding eigenvector. The normalized version of the eigenvector is then the membership function of the fuzzy set we considered when doing all pairwise assessments of the elements of its universe of discourse. The pairwise evaluations are far more convenient and manageable in comparison to any effort we make when assigning membership grades to all elements of the universe in a single step. Practically, the pairwise comparison helps the expert focus only on two elements once at a time thus reducing uncertainty and hesitation while leading to the higher level of consistency. The assessments are not free of bias and could exhibit some inconsistent evaluations. In particular, we cannot expect that the transitivity requirement could be fully satisfied. Fortunately, the lack of consistency could be quantified and monitored. The largest eigenvalue computed for R is always greater than the dimensionality of the reciprocal matrix (recall that in reciprocal matrices the elements positioned symmetrically along the main diagonal are inverse of each other), $\lambda_{max} > n$ where the equality $\lambda_{max} = n$ occurs only if the results are fully consistent. The ratio

$$\phi = (\lambda_{max} - n)/(n-1) \tag{7}$$

can be treated as an index of inconsistency of the data; the higher its value, the less consistent are the collected experimental results. This expression can be sought as the indicator of the quality of the pairwise assessments provided by the expert. If the value of ϕ is too high exceeding a certain superimposed threshold, the experiment may need to be repeated. Typically if the value of ϕ is less than 0.1 the assessment is sought to be consistent while higher values of ϕ call for the re-examination of the experimental data and a re-run of the experiment.

5.5. Fuzzy sets as granular representatives of numeric data

In general, a fuzzy set is reflective of numeric data that are put together in some context. Using its membership function we attempt to embrace them in a concise manner. The development of the fuzzy set is supported by the following experiment-driven and intuitively appealing rationale:

(a) first, we expect that A reflects (or matches) the available experimental data to the highest extent, and

(b) second, the fuzzy set is kept specific enough so that it comes with a well-defined semantics.

These two requirements points at the multiobjective nature of the construct: we want to maximize the coverage of experimental data (as articulated by (a)) and minimize the spread of the fuzzy set (as captured by (b)). These two requirements give rise to a certain optimization problem. Furthermore, which is quite legitimate, we assume that the fuzzy set to be constructed has a unimodal membership function or its maximal membership grades occupy a contiguous region in the universe of discourse in which this fuzzy set has been defined. This helps us build a membership function separately for its rising and declining sections. The core of the fuzzy set is determined first. Next, assuming the simplest scenario when using the linear type of membership functions, the essence of the optimization problem boils down to the rotation of the linear section of the membership function around the upper point of the core of A; for the illustration refer to Figure 9. The point of rotation of the linear segment of this membership function is marked by an empty circle. By rotating this segment, we intend to maximize (a) and minimize (b).

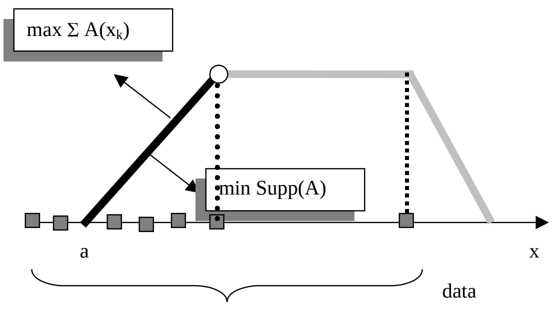


Figure 9. Optimization of the linear increasing section of the membership function of A; highlighted are the positions of the membership function originating from the realization of the two conflicting criteria.

Before moving on with the determination of the membership function, we concentrate on the location of its numeric representative. Typically, one could view an average of the experimental data $x_1, x_2, \ldots, x_n$ to be its sound representative. While its usage is quite common in practice, a better representative of the numeric data is a median value. There is a reason behind this choice. The median is a robust statistic meaning that it allows for a high level of tolerance to potential noise existing in the data. Its important ability is to ignore outliers. Given that the fuzzy set is sought to be a granular and "stable" representation of the numeric data, our interest is in the robust development not being affected by noise. Undoubtedly, the use of the median is a good starting point. Let us recall that the median is an order statistic and is formed on a basis of an ordered set of numeric values. In the case of the odd number of data in the data set, the point located in the middle of this ordered sequence is the median. When we encounter an even number of data in the granulation window, instead of picking up an average of the two points located in the middle, we consider these two points to form a core of the fuzzy set. Thus depending upon the number of data points, we either end up with triangular or trapezoidal membership function.

Having fixed the modal value of A (that could be a single numeric value, "m" or a certain interval [m, n]), the optimization of the spreads of the linear portions of the membership functions are carried out separately for their increasing and decreasing portions. We consider the increasing part of the membership function (the decreasing part is handled in an analogous manner). Referring to Figure 9, the two requirements guiding the design of the fuzzy set are and transformed into the corresponding multiobjective optimization problem as outlined as follows

(a) maximize the experimental evidence of the fuzzy set; this implies that we tend to "cover" as many numeric data as possible, viz. the coverage has to be made as high as possible. Graphically, in the optimization of this requirement, we rotate the linear segment up (clockwise) as illustrated in Figure 9. Formally, the sum of the membership grades $A(x_k)$, $\sum_k A(x_k)$ where A is the linear membership function to be optimized and x_k is located to the left to the modal value) has to be maximized

(b) Simultaneously, we would like to make the fuzzy set as specific as possible so that is comes with some well defined semantics. This requirement is met by making the support of A as small as possible, that is $\min_a|m - a|$

To accommodate the two conflicting requirements, we combine (a) and (b) in the form of the ratio that is maximized with respect to the unknown parameter of the linear section of the membership function

$$\max_a \frac{\sum_k A(x_k)}{|m-a|} \tag{8}$$

The linearly decreasing portion of the membership function is optimized in the same manner. The overall optimization returns the parameters of the fuzzy number in the form of the lower and upper bound (a and b, respectively) and its support (m or [m,n]). We can write down such fuzzy numbers as A(a, m, n, b). We exclude a trivial solution of a =m in which case the fuzzy set collapses to a single numeric entity.

As an illustration, let us consider a scenario where experimental numeric data are governed by some uniform probability density function defined over the range [0, b], b >0 that is p(x) =1/b over the [0, b] and 0 otherwise. The linear membership function of A is the one of the form A(x) = max (0, 1-x/a). The modal value of A is equal to zero. The optimization criterion (12) now reads as

$$V(a) = \frac{\int_0^a A(x)p(x)dx}{a} = \frac{1}{ab}\int_0^a (1-\frac{x}{a})dx = \frac{1}{ab}(b - \frac{b^2}{2a}) = \frac{2a-b}{2a^2} \tag{9}$$

The plot of V regarded as a function of the optimized slope of A is shown in Figure 10; here the values of "b" were varied to visualize an effect of this parameter on the behavior of V.

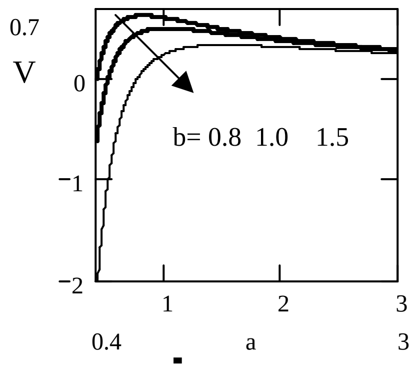


Figure 10. Plots of V versus "a" for selected values of "b".

The optimal value of "a" results from the relationship dV/da =0 and this leads to the equality a = b. The form of the relationship V= V(a) is highly asymmetric; while the values of "a" higher than the optimal value (a^{opt}) leads to a very slow degradation of the performance (V changes slowly), the rapid changes in V are noted for the values of "a" which are lower than the optimal value.

5.6. From numeric data to fuzzy sets: the essence of fuzzy clustering

Fuzzy sets can be formed on a basis of numeric data through their clustering (groupings). The groups of data give rise to membership functions that convey a global more abstract view at the available data. With this regard Fuzzy C-Means (FCM, for brief) is one of the commonly used mechanisms of fuzzy clustering (Bezdek, 1981; Pedrycz, 2005).

Let us review its formulation, develop the algorithm and highlight the main properties of the fuzzy clusters. Given a collection of n-dimensional data set $\{\mathbf{x}_k\}$, k=1,2,…,N, the task of determining its structure – a collection of "c" clusters, is expressed as a minimization of the following objective function (performance index) Q being regarded as a sum of the squared distances

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{m} \| \mathbf{x}_k - \mathbf{v}_i \|^2 \qquad (10)$$

where $\mathbf{v}_i$ s are n-dimensional prototypes of the clusters, i=1, 2,..,c and U = $[u_{ik}]$ stands for a partition matrix expressing a way of allocation of the data to the corresponding clusters; u_{ik} is the membership degree of data $\mathbf{x}_k$ in the i-th cluster. The distance between the data $\mathbf{z}_k$ and prototype $\mathbf{v}_i$ is denoted by ||.||. The fuzzification coefficient m (>1.0) expresses the impact of the membership grades on the individual clusters.

A partition matrix satisfies two important properties

$$\text{(a)}\ 0 < \sum_{k=1}^{N} u_{ik} < N, \quad i = 1,2,..., c$$

$$\text{(b)}\ \sum_{i=1}^{c} u_{ik} = 1, \quad k = 1,2,..., N \qquad (11)$$

Let us denote by **U** a family of matrices satisfying the conditions (a)-(b). The first requirement means that each cluster has to be nonempty and different from the entire set. The second requirement states that the sum of the membership grades should be confined to 1.

The minimization of Q completed with respect to $U \in \mathbf{U}$ and the prototypes $\mathbf{v}_i$ of $V=\{\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_c\}$ of the clusters. More explicitly, we write it down as follows

$$\min Q \quad \text{with respect to } U \in \mathbf{U}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c \in \mathbf{R}^n$$

The fuzzification coefficient exhibits a direct impact on the geometry of fuzzy sets generated by the algorithm. Typically, the value of "m" is assumed to be equal to 2.0. Lower values of m (that are closer to 1) yield membership functions that start resembling characteristic functions of sets; most of the membership values become localized around 1 or 0. The increase of the fuzzification coefficient (m = 3, 4, etc.) produces "spiky" membership functions with the membership grades equal to 1 at the prototypes and a fast decline of the values when moving away from the prototypes. In addition to the varying shape of the membership functions, observe that the requirement put on the sum of membership grades imposed on the fuzzy sets yields some rippling effect: the membership functions are not unimodal but may exhibit some ripples whose intensity depends upon the distribution of the prototypes and the values of the fuzzification coefficient.

While the number of clusters is typically limited to a few information granules, we can easily proceed with successive refinements of fuzzy sets. This can be done by splitting fuzzy clusters of the highest heterogeneity (Pedrycz and Reformat, 2006). Let us assume that we have already constructed "c" fuzzy clusters. Each of them can be characterized by the performance index

$$V_i = \sum_{k=1}^{N} u_{ik}^m \| \mathbf{x}_k - \mathbf{v}_i \|^2 \tag{12}$$

i = 1,2, …, c. The higher the value of V_i, the more heterogeneous the i-th cluster. The one with the highest value of V_i, that is the one for which we have $i_0 = \arg\max_i V_i$ is refined by being split into two clusters. Denote the set of data associated with the i_0-th cluster by $\mathbf{X}(i_0)$,

$$\mathbf{X}(i_0) = \{ x_k \in \mathbf{X} | \; u_{i_0 k} = \max_i u_{ik} \} \tag{13}$$

We cluster the elements in $\mathbf{X}(i_0)$ by forming two clusters which leads to two more specific (detailed) fuzzy sets. This gives rise to a hierarchical structure of the family of fuzzy sets as illustrated in Figure 11. The relevance of this construct in the setting of fuzzy sets is that it emphasizes the essence of forming a hierarchy of fuzzy sets rather than working with a single level structure of a large number of components whose semantics could not be retained.

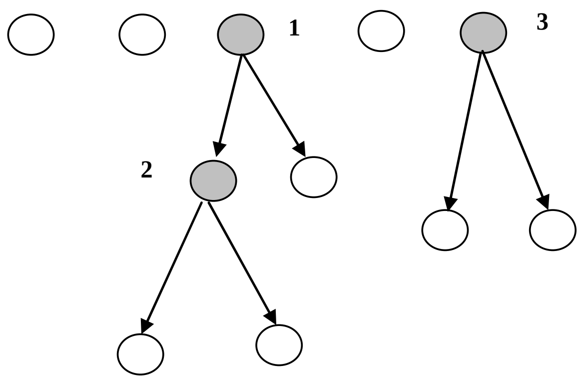


Figure 11. Successive refinements of fuzzy sets through fuzzy clustering applied to the clusters of the highest heterogeneity. The numbers indicate the order of the splits.

The process of further refinements is realized in the same by picking up the cluster of the highest heterogeneity and its split into two consecutive clusters. It is worth emphasizing that the FCM algorithm is a highly representative method of membership estimation that profoundly dwells on the use of experimental data. In contrast to some other techniques presented so far that are also data-driven, FCM can easily cope with multivariable experimental data.

6. Communication mechanisms in Granular Computing

Granular Computing can be realized in various formal frameworks. Different phenomenon can be captured by various models developed in terms of the given formalism of information granules. For instance, we may refer here to fuzzy models, rough models, and interval models. If we intend to develop an interaction between such models, they have to communicate their findings in a way the results are "understood" by the other models. There are two fundamental dimensions of the communication processes as illustrated in Figure 12. The first one is concerned with the level of granularity of information granules. The second one deals with the formalisms of Granular Computing.

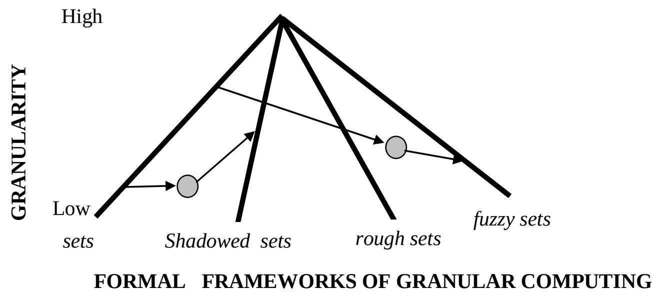


Figure 12. A two-dimensional communication plane of Granular Computing. Note that that at the highest level of granularity (which concerns numeric entities) all formalisms of information granules coincide

The role of communication mechanisms is to facilitate an interaction between the various constructs shown in the two-dimensional plane. Small circles shown in this figure denote the pertinent communication modules. For instance, consider that the construct of fuzzy sets is going to be communicated to the framework of interval analysis. This requires that the information granules (viz. fuzzy sets) are converted into a certain set. In this case the transformation (communication) mechanism is well-known: any fuzzy set can be approximated by a certain α-cut (set). The choice of the threshold level (α) itself can be optimized.

7. Rule-based systems as granular models

Granular models, as the name stipulates, are modeling constructs that are built at the level of information granules. Mappings between the granules express the relationships captured by such models. The granularity of information that is explicitly inbuilt into the construct offers interesting and useful features of the model including its evident transparency and flexibility. The same phenomenon can be viewed from different perspectives which could be highly diversified as far as the level of detail captured by the model is concerned. Similarly, we can envision a need for some interaction between the models formed at the distinct levels of granularity.

Fuzzy rule-based systems (models) are typical and commonly encountered examples of granular models. These systems are highly modular and easily expandable fuzzy models composed of a family of conditional "if – then" statements (rules) where fuzzy sets occur in their conditions and conclusions. In general, we may talk about rules embracing information granules expressed in any other formalism. The standard format of the rule with many inputs (conditions) arises in the form

-if condition_1 is A *and* condition_2 is B *and … and* condition_n is W then conclusion is Z (14)

where A, B, C,…W, Z are fuzzy sets defined in the corresponding input and output spaces. The models support a principle of locality and a distributed nature of modeling as each rule can be interpreted as an individual local descriptor of the data (problem) which is invoked by the fuzzy sets defined in the space of conditions (inputs). The local nature of the rule is directly expressed through the support of the corresponding fuzzy sets standing in its condition part. The level of generality of the rule depends upon many aspects that could be easily adjusted making use of the available design components associated with the rules. In particular, we could consider fuzzy sets of condition and conclusion whose granularity could be adjusted so that we could easily capture the specificity of the problem. By making the fuzzy sets in the condition part very specific (that is being of high granularity) we come up with the rule that is very limited and confined to some small region in the input space. When the granularity of fuzzy sets in the condition part is decreased, the generality of the rule increases. In this way the rule could be applied to more situations. To emphasize a broad spectrum of possibilities emerging in this way, refer to Figure 13 which underlines the very nature of the cases discussed above.

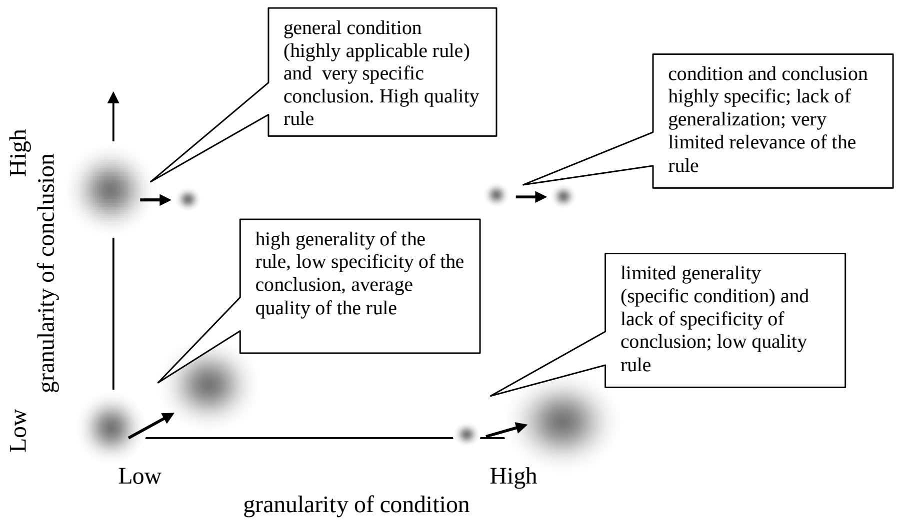


Figure 13. Examples of rules and their characterization with respect to the level of granularity of condition and conclusion parts

Quite often, we can envision a case when granular models are used in a numeric mode meaning that we are concerned with the issue of numeric outcomes produced by the models. This implies a numeric way of the optimization of the models; in these cases we encounter a minimization problem in which a numeric manifestation of the granular model is compared with the numeric target value. In other words, in the development of the granular mode we follow the scheme shown in Figure 14. More formally, we minimize the following performance index

$$Q = \sum_{k=1}^{N} \| t_k - D(Y_k)) \|^2 \qquad (15)$$

where $Y_k = G(\mathbf{x}_k)$ is the output of the granular model (G) while D denotes a decoding process (transforming the granular output into a single numeric output). $\mathbf{x}_k$ and t_k are the input-output pairs of data, k=1,2, …, N being used in the development of the model. ||.|| stands for some distance function.

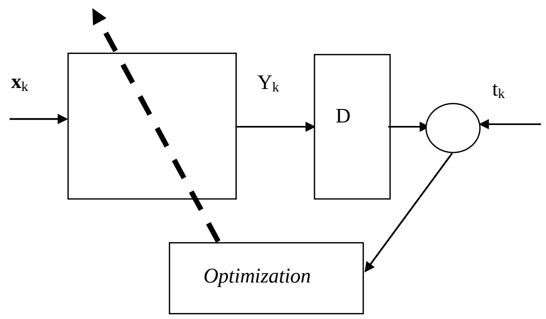


Figure 14. The optimization of granular models; note a role of the decoding module which decodes output information granule into a single numeric entity

This optimization does not allow taking into consideration the granularity of the model. An alternative approach would be to augment the performance index so that the granularity of the granular output of the model is taken into consideration. An example of such performance index could be composed as follows

$$Q = \sum_{k=1}^{N} (1 - Y_k(t_k))\Phi(Y_k) \quad (16)$$

The first term expresses a degree of compatibility of t_k and the granular output of the model (denoted by Y_k). The second one, denoted here by $F(Y_k)$ deals with the granularity of Y_k. In particular, one could consider here any energy measure of the information of the information granule (say, support of set, fuzzy set or alike). Our objective is to miminize (16) by choosing suitable values of the parameters of the granular model.

8. Conclusions

In this study, we reviewed the fundamentals of Granular Computing by stressing the role of this paradigm in the development of intelligent systems. The inherent human centricity of such systems makes the processing carried out at the level of information granules to become their integral feature. We showed a significant diversity of the underlying formalisms of information granules (including fuzzy sets, rough sets, shadowed sets, and interval analysis) and demonstrated how Granular Computing forms a unified and a highly coherent view at these mechanisms. Several ways of forming information granule were also presented. We stressed a need for effective communication mechanisms in Granular Computing where these mechanisms have to deal with various levels of granularity of information as well as various formal schemes of representation of the granules. We also underlined an issue of dealing with granular information in system modeling where the models themselves are granular constructs.

The paper serves as a brief introduction to the emerging discipline and does not pretend to cover a wealth of its conceptual developments, algorithmic pursuits and applications. It rather brings several ideas that have been around for some time, stresses the coherency of the area and emphasizes several key challenges (including communication schemes) lying ahead.

References

[1] A. Bargiela, W. Pedrycz, Granular Computing: An Introduction, Kluwer Academic Publishers, Dordercht, 2003.
[2] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, N. York, 1981.
[3] G. Cantor, Grundlagen einer Allgemeinen Mannigfaltigkeitslehre, Teubner, Leipzig, 1883.
[4] G. Cantor, Beitraage zur Begraundung der transfniten Mengenlehre. Mathematische Annalen, 46, 1895, 207-246.
[5] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, Int. J. General Systems, 17, 2–3, 1990, 191–209.
[6] E. Frias-Martinez et al., Modeling human behavior in user-adaptive systems: recent advances using soft computing techniques, Expert Systems with Applications, 29, 2005, 320-329.
[7] S. Gottwald, Mathematical fuzzy logic as a tool for the treatment of vague information, Information Sciences, 172, 1-2, 2005, 41-71.
[8] E. Hansen, A generalized interval arithmetic, Lecture Notes in Computer Science, Springer Verlag, vol. 29, 1975, 7-18.
[9] L. Jaulin, M. Kieffer, O. Didrit, E. Walter, Applied Interval Analysis, Springer, London, 2001.
[10] G.J. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995.
[11] T.Y. Lin, Data mining and machine oriented modeling: a granular computing approach, J. of Applied Intelligence, 13, 2, 2000, 113-124.
[12] J. Łukasiewicz, O logice trójwartościowej, Ruch Filozoficzny, 5, 1920, 170.
[13] J. Łukasiewicz, Philosophische Bemerkungen zu mehrwertigen Systemen des Aussagenkalk, C. R. Soc. Sci. Lettres de Varsovie, 23, 1930, 51-77.
[14] J. Łukasiewicz, Selected Works, edited by L. Borkowski. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1970.
[15] R. Moore, Interval Analysis, Prentice Hall, Englewood Cliffs, NJ, 1966.
[16] Z. Pawlak, Rough sets, Int. J. Comput. Inform. Sci. 11, 1982, 341–356.
[17] Z. Pawlak, Rough Sets. Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Dordercht, 1991.
[18] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences, 177, 1, 2007, 3-27.
[19] Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences, 177, 1, 2007, 28-40.
[20] Z. Pawlak, A. Skowron, Rough sets and Boolean reasoning, Information Sciences, 177, 1, 2007, 41-73.
[21] A. Pedrycz, M. Reformat, Hierarchical FCM in a stepwise discovery of structure in data, Soft Computing, 10, 2006, 244-256.
[22] W. Pedrycz, Computational Intelligence: An Introduction. CRC Press, Boca Raton, Fl, 1997.
[23] W. Pedrycz, Shadowed sets: representing and processing fuzzy sets, IEEE Trans. on Systems, Man, and Cybernetics, part B, 28, 1998, 103-109.
[24] W. Pedrycz (ed.), Granular Computing: An Emerging Paradigm, Physica-Verlag, Heidelberg, 2001.
[25] W. Pedrycz, Knowledge-based Clustering, J. Wiley, Hoboken, NJ, 2005.
[26] W. Pedrycz, F., Gomide, Fuzzy Systems Engineering: Toward Human-Centric Computing, J. Wiley, Hoboken, NJ, 2007.
[27] L. Polkowski, A. Skowron (eds.), Rough Sets in Knowledge Discovery, Physica-Verlag, Heidelberg, 1998.
[28] Skowron, Rough decision problems in information systems, Bulletin de l'Academie Polonaise des Sciences (Tech), 37, 1989, 59-66.
[29] M. Warmus, Calculus of approximations, Bulletin de l'Academie Polonaise des Sciences, 4, 5, 1956, 253-259.
[30] Vasilakos, W. Pedrycz (eds.), Ambient Intelligence, Wireless Networking, and Ubiquitous Computing, Artech House, Boston, MA, 2006.
[31] L.A. Zadeh, Fuzzy sets, Information & Control, 8, 1965, 338-353.
[32] L.A. Zadeh, Fuzzy logic = Computing with words, IEEE Trans. on Fuzzy Systems, 4, 1996, 103-111.
[33] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems, 90, 1997, 111-117.
[34] L.A. Zadeh, From computing with numbers to computing with words-from manipulation of measurements to manipulation of perceptions, IEEE Trans. on Circuits and Systems, 45, 1999, 105-119.
[35] L. A. Zadeh, Toward a generalized theory of uncertainty (GTU)—an outline, Information Sciences, 172, 2005, 1- 40.
[36] H.J. Zimmermann. Fuzzy Set Theory and Its Applications, 3rd ed., Kluwer Academic Publishers, Norwell, MA, 1996.