

Human centricity in computing with fuzzy sets: an interpretability quest for higher order granular constructs

Witold Pedrycz

Received: 19 August 2009 / Accepted: 19 August 2009 / Published online: 27 November 2009
© Springer-Verlag 2009

Abstract The intent of this study is to investigate the capabilities of granular computing (and computing with fuzzy sets, in particular) that are available in the currently existing framework to support the design of human-centric systems. Being cognizant of the inherently numeric nature of fuzzy sets (membership functions), we propose several essential enhancements in order to further foster their interpretability capabilities. Type-2 fuzzy sets are discussed in this setting. It is shown that they capture individual fuzzy sets and the result of their aggregation (along with the ensuing diversity) becomes reflected in the higher type of the fuzzy set. Type-2 fuzzy sets can emerge as a result of linguistic interpretation of the original numeric membership grades. The study brings forward a detailed algorithmic framework leading to the determination of type-2 fuzzy sets: in the case of aggregation, the principle of justifiable granularity is a computational vehicle while in case of linguistic interpretation we introduce a certain optimization scheme minimizing entropy which associates with the interpretation of membership functions through a limited codebook of linguistic labels. The data-driven aspects of information granules are also discussed; here we elaborate on some constructs, which rely on domain knowledge as well as experimental (predominantly numeric) evidence; this concerns both knowledge-based clustering and statistically-inclined logic operators.

Keywords Fuzzy sets of higher type · Principle of justifiable granularity · Linguistic membership · Interpretability · Statistically grounded logic operators · Knowledge-based clustering

1 Introductory notes

Ambient intelligence (AmI) and humanized computing (Cook et al. 2009) or human-centric computing (Pedrycz and Gomide 2007), in general, are the paradigms which allow humans interact with computing environment in a seamless manner, cf. also Benini et al. (2006), Kacprzyk and Zadrozny (2005), Nijholt (2008), and Remagnino and Foresti (2009). There has been a wealth of research and ensuing case studies exploring different conceptual and algorithmic avenues so that the user-system communication becomes highly intuitive, effective and responsive to the ever changing and highly dynamic needs of the users or a group of users with highly diversified needs.

Information granules (Pedrycz and Gomide 2007; Zadeh 2005) constitute a backbone or a conceptual skeleton of a variety of ways we perceive the world, communicate our findings and revise/update and organize our knowledge. Processing information granules is accomplished by the associated algorithmic platform that is pertinent to the formalism of granulation being used. By its very nature, the perception changes continuously, becomes highly context-dependent and may vary significantly throughout even a small and relatively homogeneous community of users. It is anticipated that the implemented communication mechanisms have to sense the changes, recognize them, and invoke further actions depending on the available evidence. Ideally, the system should become cognizant of its limits of understanding the intentions of the user

W. Pedrycz (✉)
Department of Electrical and Computer Engineering,
University of Alberta, Edmonton, AB T6R 2G4, Canada
e-mail: pedrycz@ee.ualberta.ca

W. Pedrycz
System Research Institute, Polish Academy of Sciences,
Warsaw, Poland

and by quantifying these limits, invoke further mechanisms of reconciliation of findings and eventually acquire more data to arrive at more justifiable decision process. For instance, in smart homes we might have cases where the intentions of the users are to be reconciled and eventually their needs could be identified or recognized to a significant extent. In intelligent agents, it is important that these agents become cognizant of their abilities to reason about their abilities to understand the surroundings and based on that engage in a sound interaction with other agents.

While granular computing (GC) including the technology of fuzzy sets is well positioned to address the challenges we highlighted above, there are a number of conceptual enhancements to its fundamentals and ensuing algorithmic practices. Bearing in mind the key challenges of communications and perceptions, we are interested in finding a constructive way in which the technology of fuzzy sets could be made more human-centric. This pertains both to the interpretability as well as the potential of fuzzy sets to cope with the diversity of the individual preferences expressed by fuzzy sets.

Learning from users is completed by taking full advantage of qualitative guidance being available in a series of training scenarios. When it comes to learning from numeric data, fuzzy sets and their underlying logic were not very much in rapport with such evidence, especially when it comes to deal with experimental findings. This calls for the use of fuzzy clustering—an algorithmic way of forming information granules. To be in line with human-centricity, we depart from numeric-guided optimization as being common in fuzzy clustering and concentrate on *knowledge-guided* fuzzy clustering (Pedrycz 2002; Pedrycz 2005a).

There is also another important computational construct of human-centric processing of fuzzy sets, which revolves around the definition of logic operators (*and* and *or*). In this case we incorporate experimental evidence which gives rise to so-called statistically inclined operators.

There are several main conceptual and design issues which are considered in an attempt to arrive at the enhanced platform of human-centric processing including: (a) knowledge capture and representation of information granules in terms of higher type constructs and their use in further decision-making, (b) interpretation of information granules via linguistically quantified assessment, (c) exploitation of knowledge-based clustering as a vehicle to design information granules, (d) construction of statistically-inclined logic operators. The structure of the paper is reflective of these main topics already outlined.

In the study, we use the standard notation encountered in fuzzy sets and GC. Likewise it is assumed some familiarity with the technology of fuzzy sets and AmI; here references such as Acampora and Loia (2008), Benini et al. (2006), and

Cook et al. (2009)) could be pertinent in the framework of our considerations. For illustrative purposes, we consider that a universe of discourse over which fuzzy sets are defined is finite, $\text{card}(\mathbf{X}) = n$. For instance, we consider a vector form of membership grades such as $[A(x_1) A(x_2) \dots A(x_n)]$.

2 Perceptions and their augmented models: emergence of fuzzy sets of higher type

Concepts (notions) being building blocks of constructs of AmI and its algorithmic underpinnings, are formally represented as fuzzy sets. Fuzzy set is a quantitative descriptor of a certain generic or compound notion. The compound notions are designed by making use of logic operators.

An individual could perceive the same concept differently over time. The differences may depend on context in which the concept becomes engaged. For instance, *comfortable* temperature in a smart home depends on the mode, physical status, and intentions of the individual. A family of fuzzy sets can conveniently capture the existing diversity. The same concept can be perceived in a different way by a group of users and again we may easily encounter a family of fuzzy sets. Our intent is to represent and quantify the variability of a family of a fuzzy set by a higher order construct such as type-2 fuzzy set (Mendel 2007; Walker and Walker 2009). The essence of the underlying construct is visualized in Fig. 1a. Another scenario in which type-2 fuzzy sets can be formed concerns a linguistic interpretation of a fuzzy set, Fig. 2b where we abandon the use of detailed and overly specific (precise) numeric membership values (which are predominantly visible in the current style of processing fuzzy sets) and start perceiving them more in terms of linguistically quantified (granulated) membership grades, say *high*, *medium* or *low* membership values. This quantification can be essential in the enhancements of interpretation capabilities of information granules.

Fuzzy sets of higher type, say type-3 fuzzy sets, can arise as a result of perception, which is formed in a hierarchical fashion. For instance, a collection of type-2 fuzzy sets (where each of them is a result of the aggregation achieved at the lower level) when being aggregated gives rise to type-3 fuzzy sets; see Fig. 2. Obviously, one can proceed with higher order construct assuming that are legitimized by the hierarchical structure of opinions captured by users.

The fundamental design questions that arise in the setting are concerned with the effective way of forming fuzzy sets of type-2. The proposed development is realized by invoking the principle of justifiable granularity. The interpretation aspects of fuzzy sets are revealed by constructing a family of fuzzy sets of some well-defined semantics which is quantified via some optimization process.

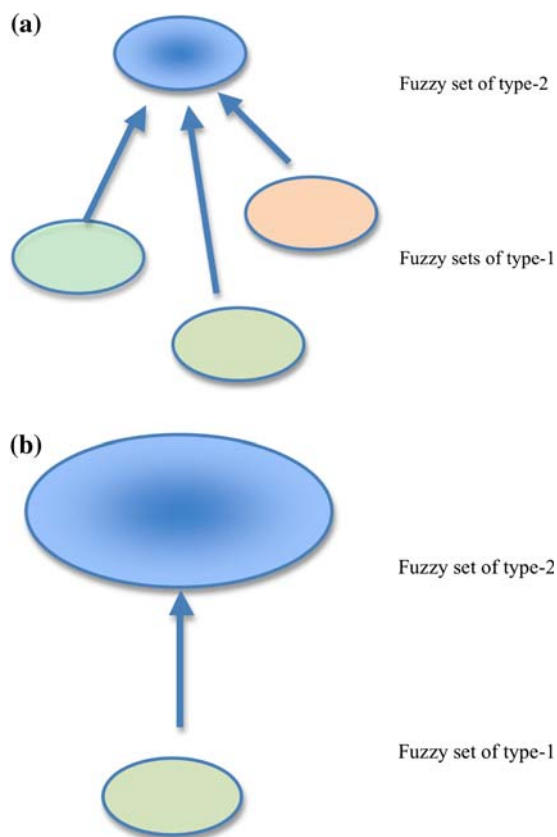


Fig. 1 Aggregation and interpretation: from a family of fuzzy sets to type-2 fuzzy set (a) and linguistically quantified fuzzy set in the form of type-2 fuzzy set (b)

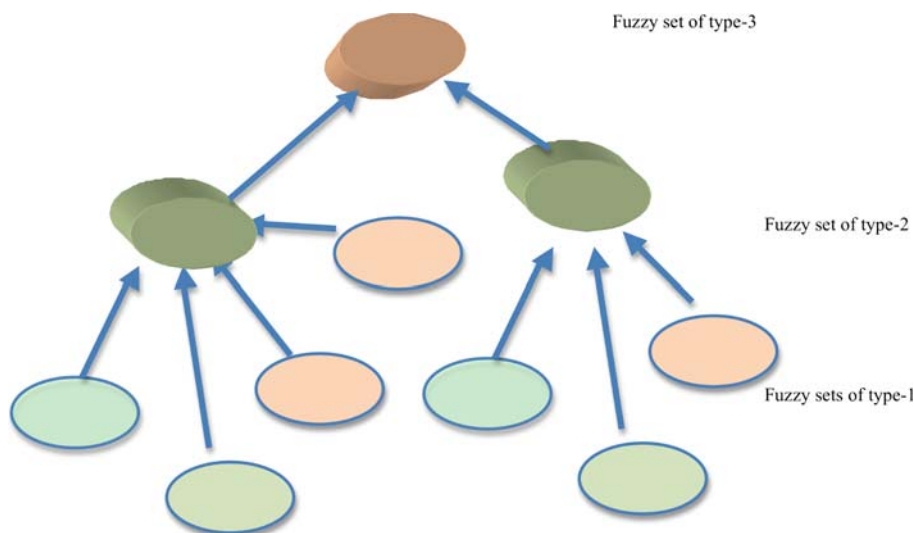
3 The design of type-2 fuzzy sets via the principle of justifiable granularity

Let us consider a family of fuzzy sets A_1, A_2, \dots, A_c defined in X and concentrate on the membership degrees

$A_1(x_j), A_2(x_j), \dots, A_c(x_j)$ for the fixed element of X . The resulting sequence of membership degrees consists of the entries (membership degrees) z_1, z_2, \dots, z_c . Denote the set of these membership values by Z . Without any loss of generality, assume that these entries are arranged in a non-decreasing order. The essence of the principle of justifiable granularity (Pedrycz and Vukovich 2002; Pedrycz 2005a) can be outlined as follows: Given a collection of information granules of a certain type, their aggregation can be represented in the form of a certain information granule of the type higher than the type of the original constructs involved in the process of aggregation. The granularity of the higher type information granule has to be as specific as possible while at the same time the information granule of the higher type should represent the individual information granules to the highest extent. These two requirements will be re-stated in a more explicit manner when moving to the detailed algorithm.

Proceeding with the algorithmic aspects, the construction of the information granule proceeds in two phases. First, we form a numeric representation of Z . There are several possible alternatives however a modified median arises as an appealing alternative. First, median is robust. Second, median is one of the elements of Z , which makes its interpretation straightforward (note the interpretation of average is far less obvious). Here we offer some slight modification of the median in the following sense. If the number of elements is odd, the median of Z is the central point, say z_0 . We have $z_0 = z_{n/2+1}$. If the number of elements in Z is even we form an interval spanned over the two central points of this sequence, $[z_{n/2-1}, z_{n/2+1}]$. Let us introduce the following notation: For odd number of elements in Z :

Fig. 2 Successive aggregation of differently perceived concepts producing fuzzy sets of higher type (type-3 in this case)



$$\mathbf{Z}_- = \{z_i | z_i < z_{n/2+1}\} \quad \mathbf{Z}_+ = \{z_i | z_i > z_{n/2+1}\} \quad (1)$$

For even number of elements in \mathbf{Z} :

$$\mathbf{Z}_- = \{z_i | z_i < z_{n/2-1}\} \quad \mathbf{Z}_+ = \{z_i | z_i > z_{n/2+1}\} \quad (2)$$

In what follows, we will be discussing the construction of information granule which is realized separately for \mathbf{Z}_- and \mathbf{Z}_+ . We also decide to form an interval-like information granule; refer to Fig. 3. As we can see, the two parameters (bounds) of the interval (a and b) are optimized individually.

The two requirements expressed in the principle of justifiable granularity are now quantified as follows:

- high coverage of data. Ideally, we would like to see the interval $[a, z_0]$ cover all data
- high specificity. The length (support) of the interval is as small as possible

As we intend to maximize the sum $\text{card}\{i | z_i \in [a, z_0]\}$ (a) and minimize the length of the interval $|z_0 - a|$, as implied by (b) these two criteria can be coming in the following performance index V whose values to be maximized.

$$V = \text{card}\{i | z_i \in [a, z_0]\} / |z_0 - a| \quad (3)$$

The optimal value of “ a ” results from this maximization. The determination of the optimal value of “ b ” is carried out in an analogous fashion. One can generalize (3) by introducing an extra parameter γ assuming positive values

$$V = [\text{card}\{i | z_i \in [a, z_0]\}]^\gamma / |z_0 - a| \quad (4)$$

In case of $\gamma < 1$ we achieve a dilution effect while the values of γ higher than one yield a concentration effect.

Realizing the above optimization for the consecutive elements of \mathbf{X} , the result becomes an interval-valued type-2 fuzzy set. The principle of justifiable granularity is general enough so that the result could be any granular construct. For instance, we can assume a certain form of membership functions and apply the same principle of balancing the

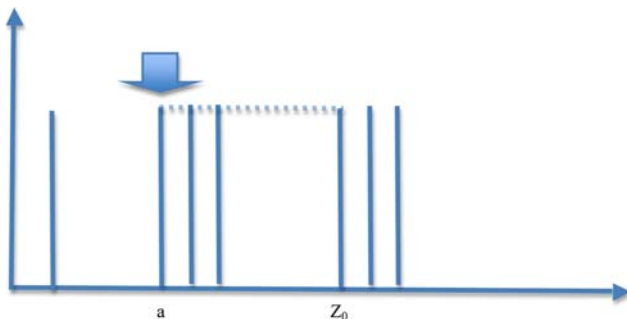


Fig. 3 The realization of the principle of justifiable granularity: satisfaction of the coverage and specificity requirements

uncertainty as before. There is some minor modification of (3) considering that we are concerned now with the membership degrees not Boolean characteristic functions. We obtain

$$V = \sum_{i=1}^c A(z_i) / |z_0 - a| \quad (5)$$

where A is a fuzzy set (membership function) to be constructed. Likewise we can also engage the optimization index of the form

$$V = \sum_{i=1}^c A^\gamma(z_i) / |z_0 - a| \quad (6)$$

The two forms of membership functions, which are typically of practical interest, deal triangular and parabolic membership functions. The functions could be asymmetric as the optimization is carried out independently for the increasing and decreasing sections of the membership functions. The Gaussian membership functions could be another option and again they could be made asymmetric by allowing for two spreads of the membership functions..5

4 Interpretability of fuzzy sets: a framework of type-2 fuzzy sets

Considering the second situation outlined in Fig. 1b, we envision a situation when membership functions are to be interpreted linguistically and this school of thought triggers the usage of type-2 fuzzy sets.

Let us stress that membership functions are numeric constructs. Detailed numeric values of membership functions are too precise and in this way difficult to interpret. Therefore it becomes advantageous to come up with a more descriptive quantification, say *low*, *medium*, *high* membership levels. In this sense, rather than talking about “a membership degree of 0.75”, we might have a far more descriptive quantification such as “*high* membership”. Figure 4 illustrates the essence of this construct. Crucial to this development are fuzzy sets of linguistic values defined in the $[0,1]$ interval of membership grades.

The construction of these fuzzy sets defined in $[0,1]$ and used in the interpretation of the original membership function is required. More formally, let us consider a fuzzy set of interest defined in some space \mathbf{X} , $A: \mathbf{X} \rightarrow [0,1]$. Consider “ r ” fuzzy sets—linguistic descriptors of membership grades $\mathbf{B} = \{B_1, B_2, \dots, B_r\}$, where $B_i: [0,1] \rightarrow [0,1]$. For instance (what is quite common in practice), B_i s can be sought as triangular membership functions. To benefit from the descriptive power supplied by the \mathbf{B} , its fuzzy sets have to be optimized where the optimization could be guided by different design criteria.

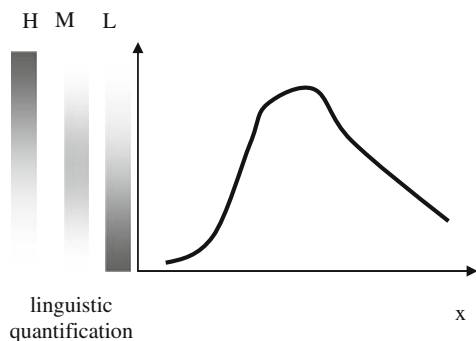


Fig. 4 Fuzzy set of type-1 and its linguistic interpretation through fuzzy sets of linguistic quantification

Here we highlight two interesting alternatives:

- (a) determination of fuzzy sets through clustering or fuzzy clustering of membership degrees. The points located in-between the prototypes (note that we are concerned with a one-dimensional case) determine the bounds of the intervals describing the elements of **B** (regarded in this case as intervals). In case of fuzzy clustering, we arrive at the fuzzy sets forming as linguistic quantifiers of the membership levels.
- (b) Construction of **B** carrying out minimal uncertainty. More specifically, we consider a representation criterion where we request that A is well-defined in terms of the elements of **B** meaning that there is the least amount of uncertainty associated with such representation. As the uncertainty can be conveniently captured through the concept of entropy, we engage it in the formulation of the performance index to be minimized. Let us compute the entropy of A expressed in terms of B_i for a certain element x of **X**, that is $H(B_i(A(x)))$. The overall entropy determined over all elements of **B** and the entire universe of discourse **X** comes in the form

$$V = \int_{\mathbf{X}} \sum_{i=1}^r H(B_i(A(x))) dx \tag{7}$$

Let us recall that the entropy function $H(z)$ defined in $[0,1]$, $H: [0,1] \rightarrow [0,1]$, satisfies well-known requirements: (a) monotonic increase in $[0, 1/2]$ and monotonic decrease in $[1/2, 1]$ and (b) boundary conditions of the form $H(0) = H(1) = 0$, $H(1/2) = 1.0$. The minimization of (7) is expressed as $\text{Min}_{\mathbf{a}} H(\mathbf{B})$ where \mathbf{a} is a vector of parameters of the fuzzy sets in **B**. In particular, when B_i s are taken as triangular fuzzy numbers with $1/2$ overlap between neighboring fuzzy sets, then the minimum of (12) is achieved by adjusting the modal values of the fuzzy sets (viz. the modal values are the corresponding coordinates of the vector **a**). Here mechanisms of evolutionary

optimization are worth pursuing. Note that the distribution of the modal values need not be uniform but it rather reflects the characteristics of the data. In case of finite space **X**, the integral standing in (7) is replaced by the corresponding sum, that is

$$V = \sum_{j=1}^n \sum_{i=1}^r H(B_i(x_j)) \tag{8}$$

Once the fuzzy sets of linguistic quantification of the fuzzy set defined in the finite space have been defined, the original membership function is translated into a string of labels, say LLMMMMHML, etc. with L, M, and H standing for the linguistic labels (say, being associated with $B_1(L)$, $B_2(M)$, and $B_3(H)$). Note that in this interpretation, any $x \in \mathbf{X}$ and subsequently $A(x)$ invokes (matches) several B_i 's to a nonzero degree. In the formation of the string of the labels, we pick up this fuzzy set B_i to which x belongs to the highest extent. The essence of the linguistic interpretation is illustrated in Fig. 5.

As an illustrative example of the concepts discussed above, we consider a unimodal fuzzy set with the membership function shown in Fig. 6.

To come up with the linguistic quantification, we run *K*-means clustering whose results produce the three intervals formed over the $[0,1]$ range, that is

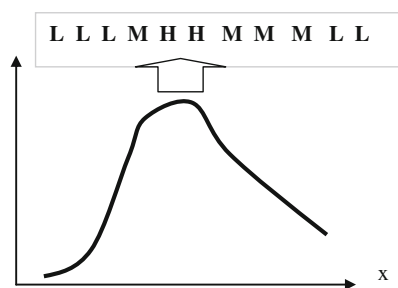


Fig. 5 From numeric membership grades to a string of linguistic labels

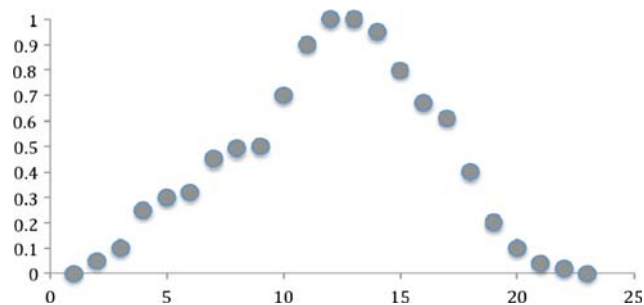


Fig. 6 Fuzzy set A with a unimodal membership function

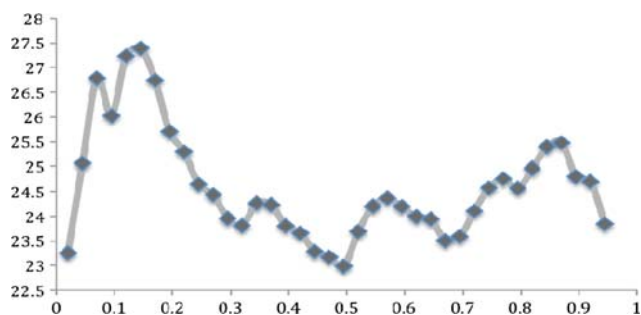


Fig. 7 Plot of V treated as a function of the modal value of the *Medium* fuzzy set of quantification of the membership function

$$S = [0.00 \quad 0.23] \quad M = [0.23 \quad 0.73] \\ L = [0.73 \quad 1.00]$$

Each of them comes with a straightforward meaning. The resulting construct here is a family of interval-valued fuzzy sets.

To move with the entropy-based criterion where we consider three triangular fuzzy sets expressed in $[0,1]$ with $\frac{1}{2}$ overlap between neighboring fuzzy sets, we consider this to be an optimization problem where the modal value of the intermediate fuzzy set (quantified as *Medium*) is subject to the minimization of the entropy criterion. The entropy function is defined in the form $h(u) = 4u(1 - u)$ (which satisfies the obvious boundary conditions of $h(0) = 0$, $h(1) = 1$, $h(1/2) = 1$). The plot of V treated as a function of the optimized modal value is illustrated in Fig. 7. The optimal position of the *Medium* quantification of the membership is equal to 0.5. It is not surprising this value is slightly different from the one reported when using the other method as the criteria being optimized are different.

One could stress here that the distribution of the obtained cutoff points depends on the membership grades.

5 Shadowed sets as a three-valued logic representation of fuzzy sets

Membership functions are numeric constructs as far as quantification of membership degrees is concerned. Linguistic quantification enhances interpretability. The concept supported by shadowed sets is developed with the similar goal in mind: rather than dealing with numeric membership degrees, we are concerned with a very limited three-valued quantification characterization scheme of membership functions.

Formally speaking, a shadowed set A (Pedrycz 1998, 2005b) defined in some space \mathbf{X} is a set-valued mapping coming in the following form

$$A : \mathbf{X} \rightarrow \{0, [0, 1], 1\} \quad (9)$$

The co-domain of A consists of three elements that is 0, 1, and the unit interval $[0,1]$. These can be treated as degrees of membership of elements to the concept captured by A . Given this simplified view, we approximate the original fuzzy with intent of enhancing the readability (interpretability) of the information granules. These three quantification levels come with an apparent interpretation. All elements of \mathbf{X} for which $A(x)$ assume 1 are called a core of the shadowed set—they embrace all elements that are fully compatible with the concept conveyed by A . The elements of \mathbf{X} for which $A(x)$ attains zero are excluded from A . The elements of \mathbf{X} for which we have assigned the unit interval are completely uncertain—we are not at position to allocate any numeric membership grade. In this region we are faced with a complete uncertainty (don't know quantification). Therefore we allow the usage of the unit interval, which reflects uncertainty meaning that any numeric value could be permitted here. In essence, such element could be excluded (we pick up the lowest possible value from the unit interval), exhibit partial membership (any number within the range from 0 and 1) or could be fully allocated to A . Given this extreme level of uncertainty (nothing is known and all values are allowed), we call these elements shadows and hence the name of the shadowed set. An illustration of the underlying concept of a shadowed set is included in Fig. 8.

One can view this mapping (viz. shadowed set) as an interesting example of a three-valued logic as encountered in the classic model introduced by Lukasiewicz. Having this in mind, we can think of shadowed sets as a *symbolic* representation of *numeric* fuzzy sets. Obviously, the elements of co-domain of A could be labeled using symbols (say, certain, shadow, excluded; or a , b , c and alike) endowed with some well-defined semantics. This conversion of a fuzzy set into a shadowed set is instrumental from the interpretation perspective. For instance, the multimodal membership function illustrated in Fig. 9, once the induced shadowed set has been constructed, comes with a three-valued logic interpretation as succinctly underlined by a string of quantifying values included in the same figure.

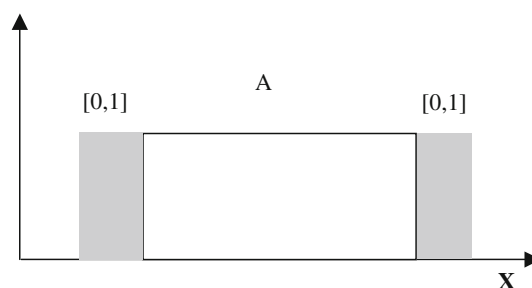
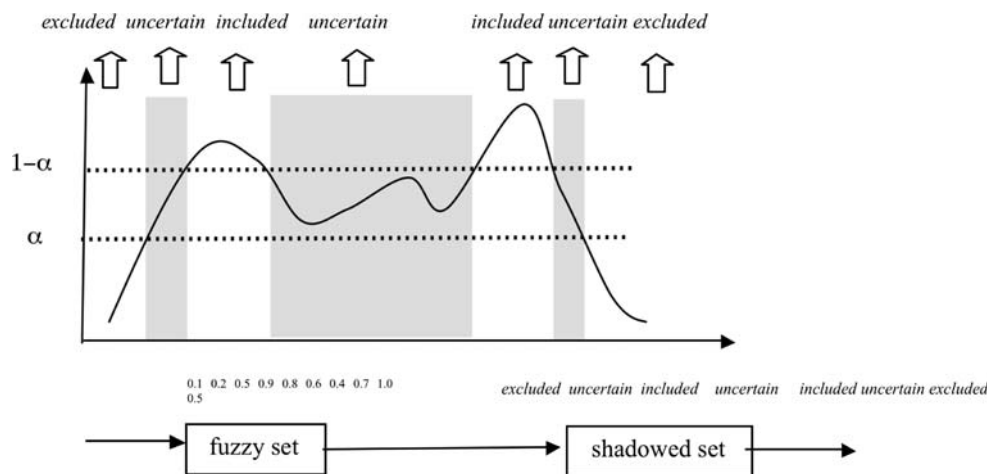


Fig. 8 An example of a shadowed set A ; note that *shadows* are formed around the cores of the construct

Fig. 9 From fuzzy set A to its shadowed set representation and the resulting interpretation



6 Statistically sound logic connectives

The statistical support incorporated into the structure of the logic connective helps address the issues emerging when dealing with the aggregation schemes and fusion of information coming from a number of sources, cf. Bouchon-Meunier (1998) and Torra (2005). These operators are also critical when being articulated by the possibility and necessity measures. We introduce a concept of statistically augmented (directed) logic connectives (Pedrycz 2009) by constructing a connective that takes into consideration a statistically driven aggregation with some weighting function being reflective of the nature of the underlying logic operation.

6.1 SOR logic connectives

The (SOR) connective is defined as follows. Denote by $w(u)$ a monotonically non-decreasing weight function from $[0,1]$ to $[0,1]$ with the boundary condition $w(1) = 1$. The result of the aggregation of the membership grades $\mathbf{z} = [z_1, z_2, \dots, z_N]$, denoted by $SOR(\mathbf{z}; w)$, is obtained as a result of the minimization of the following expression (performance index) Q

$$Q = \sum_{i=1}^N w(z_i) |z_i - y| \quad \text{Min}_y Q \tag{10}$$

where the value of “ y ” minimizing the above expression is taken as the result of the operation $SOR(\mathbf{z}, w) = y$. Writing it differently we come up with the following expression $SOR(\mathbf{z}, w) = \arg \min_{y \in [0,1]} \sum_{k=1}^N w(z_k) |z_k - y|$. The weight function “ w ” is used to model a contribution of different membership grades to the result of the aggregation. Several models of the relationships “ w ” are of particular interest; all of them are reflective of the *or* type of aggregation

(a) $w(z)$ assumes a form of a certain step function

$$w(z) = \begin{cases} 1 & \text{if } z \geq z_{\max} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where z_{\max} is the maximal value reported in \mathbf{z} . This weight function effectively eliminates all the membership grades but the largest one. For this form of the weight function, we effectively end up with the maximum operator, $SOR(\mathbf{z}, w) = \max(z_1, z_2, \dots, z_N)$

(b) $w(z)$ is equal identically to 1, $w(z) = 1$. It becomes obvious that the result of the minimization of the following expression

$$\sum_{i=1}^N |z_i - y| \tag{12}$$

is a median of \mathbf{z} , $\text{median}(\mathbf{z})$. Subsequently $SOR(\mathbf{z}, w) = \text{median}(\mathbf{z})$. Interestingly, the result of the aggregation is a robust statistics of the membership grades involved in this operation.

We can consider different forms of weight functions. In particular, one could think of an identity function $w(z) = z$. There is an interesting and logically justified alternative which links the weight functions with the logic operator standing behind the logic operations. In essence, the weight function can be induced by various *t*-conorms (*s*-norms) by defining $w(z)$ to be in the form $w(z) = zsz$. In particular, for the maximum operator, we obtain the identity weight function $w(z) = \max(z, z) = z$. For the probabilistic sum, we obtain $w(z) = (z + z - z * z) = 2z(1 - z)$. For the Lukasiewicz or connective, the weight function comes in the form of some piecewise linear relationship with some saturation region, that is

$$w(z) = \max(1, z + z) = \max(1, 2z).$$

In general, the weight functions (which are monotonically non-decreasing and satisfy the condition $w(1) = 1$) occupy the region of the unit square. For all

these weight functions implied by t -conorms, the following inequality holds: $\text{median}(\mathbf{z}) \leq \text{SOR}(\mathbf{z}, w) \leq \max(\mathbf{z})$.

6.2 SAND logic connectives

The statistically grounded AND (SAND) logic connective is defined in an analogous way as it was proposed in the development of the SOR. Here $w(z)$ denotes a monotonically non-increasing weight function from $[0,1]$ to $[0,1]$ with the boundary condition $w(0) = 1$. The result of the aggregation of $\mathbf{z} = [z_1, z_2, \dots, z_N]$, denoted by $\text{SAND}(\mathbf{z}; w)$, is obtained from the minimization of the same expression (8) as introduced before. Thus we produce the logic operator $\text{SAND}(\mathbf{z}, w) = y$ with “ y ” being the solution to the corresponding minimization problem.

As before, we can envision several models of the weight function; all of them are reflective of the *and* type of aggregation

- (a) $w(z)$ assumes a form of some step function

$$w(z) = \begin{cases} 1 & \text{if } z \leq z_{\min} \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

where z_{\min} is the minimal value in \mathbf{z} . This weight function eliminates all the membership grades but the smallest one. For this form of the weight function, we effectively end up with the maximum operator, $\text{SAND}(\mathbf{z}, w) = \min(z_1, z_2, \dots, z_N)$

- (b) for $w(z)$ being equal identically to 1, $w(z) = 1$, SAND becomes a median, namely $\text{SAND}(\mathbf{z}, w) = \text{med}(\mathbf{z})$.
- (c) more generally, the weight function is defined on a basis of some t -norm as follows, $w(z) = 1 - tz$. Depending upon the specific t -norm, we arrive at different forms of the mapping. For the minimum operator, $w(z) = 1 - \min(z, z) = 1 - z$ which is a complement of “ z ”. The use of the product operation leads to the expression $w(z) = 1 - z^2$. In the case of the Lukasiewicz *and* connective, one has $w(z) = 1 - \max(0, z + z - 1) = 1 - \max(0, 2z - 1)$.

Investigating the fundamental properties of the logic connectives, we note that the commutativity and monotonicity properties hold. The boundary condition does not hold when being considered with respect to a single membership grade (which is completely understood given the fact that the operation is expressed by taking into consideration a collection of membership grades). Assuming the t -norm and t -conorm driven format of the weight function (where we have $w(1) = 1$ and $w(0) = 0$ for *or* operators and $w(0) = 1$ and $w(1) = 1$ for *and* operators) we have $\text{SOR}(\mathbf{1}, w) = 0$, $\text{SAND}(\mathbf{0}, w) = 0$. The property of associativity does not hold. This is fully justified given that the proposed operators are inherently

associated with the overall processing of all membership grades not just individual membership values.

The possibility and necessity measures determined for the two information granules A and B being articulated in the language of SAR and SAND are expressed in the following manner

$$\begin{aligned} \text{Poss}(A, B) &= \text{SOR}(\mathbf{z}, \mathbf{w}); = a_i b_i \\ \text{Nec}(A, B) &= \text{SAND}(\mathbf{z}, \mathbf{w}); z_i = a_i b_i \end{aligned} \quad (14)$$

7 Knowledge-based clustering

In its generic version, fuzzy clustering involves data-driven optimization (Bezdek 1981). Numeric data are processed giving rise to fuzzy sets. On the other hand, knowledge-based clustering takes into account some domain knowledge hints that have to be prudently incorporated into the generic clustering procedure. Knowledge hints can be conveniently captured and formalized in terms of fuzzy sets. Altogether with the underlying clustering algorithms, they give rise to the concept of knowledge-based clustering—a unified framework in which data and knowledge are processed together in a uniform fashion.

We discuss some of the typical design scenarios of knowledge-based clustering and show how the domain knowledge can be effectively incorporated into the fabric of the original predominantly data-driven clustering techniques.

We can distinguish several interesting and practically viable ways in which domain knowledge is taken into consideration:

A subset of labeled patterns. The knowledge hints are provided in the form of a small subset of labeled patterns (data) $\mathbf{K} \subset N$ (Pedrycz 2005a). For each of them we have a vector of membership grades f_k , $k \in \mathbf{K}$ which consists of degrees of membership the pattern is assigned to the corresponding clusters. As usual, we have $f_i \in [0, 1]$ and $\sum_{i=1}^c f_{ik} = 1$.

Proximity-based clustering. Here we are provided a collection of pairs of data with their specified levels of closeness (resemblance) which are quantified in terms of proximity, $\text{prox}(k, l)$ expressed for \mathbf{x}_k and \mathbf{x}_l . The proximity offers a very general quantification scheme of resemblance: we require reflexivity and symmetry, that is $\text{prox}(k, k) = 1$ and $\text{prox}(k, l) = \text{prox}(l, k)$ however no transitivity is needed.

“belong” and “not-belong” Boolean relationships between patterns. These two Boolean relationships stress that two patterns should belong to the same clusters, $R(\mathbf{x}_k, \mathbf{x}_l) = 1$ or they should be placed apart in two different clusters, $R(\mathbf{x}_k, \mathbf{x}_l) = 0$. These two requirements could be relaxed by requiring that these two relationships return values close to one or zero.

Uncertainty of labeling/allocation of patterns. We may consider that some patterns are “easy” to assign to clusters while some others are inherently difficult to deal with meaning that their cluster allocation is associated with a significant level of uncertainty. Let $\Phi(\mathbf{x}_k)$ stands for the uncertainty measure (e.g., entropy) for \mathbf{x}_k (as a matter of fact, Φ is computed for the membership degrees of \mathbf{x}_k that is $\Phi(\mathbf{u}_k)$ with \mathbf{u}_k being the k -th column of the partition matrix. The uncertainty hint is quantified by values close to 0 or 1 depending upon what uncertainty level a given pattern is coming from.

Depending on the character of the knowledge hints, the original clustering algorithm needs to be properly refined. In particular the underlying objective function has to be augmented to capture the knowledge-based requirements. Below shown are several examples of the extended objective functions dealing with the knowledge hints introduced above.

When dealing with some labeled patterns we consider the following augmented objective function

$$Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 + \alpha \sum_{i=1}^c \sum_{k=1}^N (u_{ik} - f_{ik} b_k)^2 \|\mathbf{x}_k - \mathbf{v}_i\|^2 \tag{15}$$

where the second term quantifies distances between the class membership of the labeled patterns and the values of the partition matrix. The positive weight factor (α) helps set up a suitable balance between the knowledge about classes already available and the structure revealed by the clustering algorithm. The Boolean variable b_k assumes values equal to 1 when the corresponding pattern has been labeled.

The proximity constraints are accommodated as a part of the optimization problem where we minimize the distances between proximity values being provided and those generated by the partition matrix $P(k_1, k_2)$

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

$$\| |\text{prox}(k_1, k_2) - P(k_1, k_2)| \| \rightarrow \text{Min } k_1, k_2 \in K \tag{16}$$

with K being a pair of patterns for which the proximity level has been provided. It can be shown that given the partition matrix the expression $\sum_{i=1}^c \min(u_{ik_1}, u_{ik_2})$ generates the corresponding proximity value.

For the uncertainty constraints, the minimization problem can be expressed as follows

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

$$\| \Phi(\mathbf{u}_k) - \gamma_k \| \rightarrow \text{Min } k \in K \tag{17}$$

where K stands for the set of patterns for which we are provided with the uncertainty values γ_k .

Undoubtedly the extended objective functions call for the optimization scheme that is more demanding as far as the calculations are concerned. In several cases we cannot modify the standard technique of Lagrange multipliers, which leads to an iterative scheme of successive updates of the partition matrix and the prototypes. In general, though, the knowledge hints give rise to a more complex objective function in which the iterative scheme cannot be useful in the determination of the partition matrix and the prototypes. Alluding to the generic FCM scheme, we observe that the calculations of the prototypes in the iterative loop are doable in case of the Euclidean distance. Even the Hamming or Tchebyshev distance brings a great deal of complexity. Likewise, the knowledge hints lead to the increased complexity: the prototypes cannot be computed in a straightforward way and one has to resort himself to more advanced optimization techniques. Evolutionary computing arises here as an appealing alternative. We may consider any of the options available there including genetic algorithms, particle swarm optimization, ant colonies, to name some of them. The general scheme can be schematically structured as follows:

- repeat {EC (prototypes); compute partition matrix U;}

8 Conclusions

Human centrality is of paramount relevance to humanized computing. The success and relevance of AmI dwells on efficient communication in which computers are often invisible to a high extent. Granular computing brings an important notion of information granules—abstractions, which help establish an algorithmic fabric of AmI. One can substantially benefit from the well-established conceptual fundamentals of fuzzy sets yet those need to be augmented to make them more human-centric. We highlighted two important facets to be developed here, that is a construction of higher type of fuzzy sets who emergence is justified by the diversity of users (implying that the concepts one has to deal with exhibit some diversity) and the linguistic quantification of fuzzy sets. In both situations, we witness an important role of type-2 fuzzy sets or, more generally, fuzzy sets of higher type. Interestingly, our investigations offer an important and highly legitimate justification behind the emergence and design of type-2 fuzzy sets; where this convincing justification was missing in the past. The principle of justifiable granularity comes here as an algorithmic vehicle used to quantify the diversity of

membership degrees (which in turn reflect the variability in concept description of a single individual or difference in concept characterization varying over a group of users). The granular descriptors are reflective of the domain knowledge as well as some numeric experimental evidence- the effective exploitation of these two important sources of knowledge is captured by fuzzy clustering enhanced by available domain knowledge (viz. knowledge-based clustering) and more careful accommodation of statistical evidence stemming from information granules being aggregated as exemplified here by statistically grounded operators.

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

References

- Acamora G, Loia V (2008) A proposal of ubiquitous fuzzy computing for ambient intelligence. *Inf Sci* 178(3):631–646
- Benini L, Farella E, Guiducci C (2006) Wireless sensor networks: enabling technology for ambient intelligence. *Microelectron J* 37(12):1639–1649
- Bezdek JC (1981) *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, New York
- Bouchon-Meunier B (1998) *Aggregation and fusion of imperfect information*. Physica-Verlag, Heidelberg
- Cook DJ, Augusto JC, Jakkula VR (2009) *Ambient intelligence: technologies, applications, and opportunities*. Pervasive Mobile Comput (in press)
- Kacprzyk J, Zadrozny S (2005) Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools. *Inf Sci* 173(4):281–304
- Mendel JM (2007) Advances in type-2 fuzzy sets and systems. *Inf Sci* 177(1):84–110
- Nijholt A (2008) Google home: experience, support and re-experience of social home activities. *Inf Sci* 178(3):612–630
- Pedrycz W (1998) Shadowed sets: representing and processing fuzzy sets. *IEEE Trans Syst Man Cybern B* 28:103–109
- Pedrycz W (2002) Collaborative fuzzy clustering. *Pattern Recogn Lett* 23:675–686
- Pedrycz W (2005a) *Knowledge-based fuzzy clustering*. Wiley, New York
- Pedrycz W (2005b) Interpretation of clusters in the framework of shadowed sets. *Pattern Recogn Lett* 26(15):2439–2449
- Pedrycz W (2009) Statistically grounded logic operators in fuzzy sets. *Eur J Oper Res* 193(2):520–529
- Pedrycz W, Gomide F (2007) *Fuzzy systems engineering*. Wiley, Hoboken
- Pedrycz W, Vukovich G (2002) Clustering in the framework of collaborative agents. *Proceedings 2002 IEEE international conference on fuzzy systems*, vol 1, pp 134–138
- Remagnino P, Foresti GL (2009) Computer vision methods for ambient intelligence. *Image Vis Comput* (in press)
- Torra V (2005) Aggregation operators and models. *Fuzzy Sets Syst* 156:407–410
- Walker CL, Walker EA (2009) Sets with type-2 operations. *Int J Approx Reason* 50(1):63–71
- Zadeh LA (2005) Toward a generalized theory of uncertainty (GTU)—an outline. *Inf Sci* 172:1–40