Interpretability of Fuzzy Information Granules

Corrado Mencar

University of Bari, Italy
mencar@di.uniba.it

Abstract. Human-Centric Information Processing requires tight communication processes between users and computers. These two actors, however, traditionally use different paradigms for representing and manipulating information. Users are more inclined in managing perceptual information, usually expressed in natural language, whilst computers are formidable number-crunching systems, capable of manipulating information expressed in precise form. Fuzzy information granules could be used as a common interface for communicating information and knowledge, because of their ability of representing perceptual information in a computer manageable form. Nonetheless, this connection could be established only if information granules are interpretable, i.e. they are semantically co-intensive with human knowledge. Interpretable information granulation opens several methodological issues, regarding the representation and manipulation of information granules, the interpretability constraints and the granulation processes. By taking into account all such issues, effective Information Processing systems could be designed with a strong Human-Centric imprint.

1 Introduction

Human-Centered Computing (HCC) is a new field embracing all the methodologies that apply to applications in which people directly interact with computer technologies. Thus, HCC refers to a modern way of tackling computing issues by taking into account user needs and constraints [1, Ch. 1].

We stress the importance of communication between users and machines, the former acting as producers/consumers of information and the latter being involved in the concrete task of information processing. Besides, we observe the different paradigms for interpreting and manipulating information by users and computers. Users are indeed more inclined in managing perceptual information, usually expressed in natural language, whilst computers are formidable number-crunching systems, capable of manipulating information expressed in precise form.

The "semantic gap" between users and machines is apparent. Quite often, this gap is filled by users, which support the effort of translating perceptual information into computer-understandable forms and interpreting computer results. This approach requires technically skilled users and prevents computers to be easily used by other people who may take full advantage from more "humanized" machines (e.g. physicians, managers, decision makers, etc.) [2].

From the last decade, however, a paradigm shift – the so-called Human Centric Computing – is in act [3]. The great enhancement of computing technologies, as well as the birth and consolidation of new computing models (e.g. Granular Computing) are encouraging the development of novel techniques and methods that enable

A. Bargiela, W. Pedrycz (Eds.): Human-Centric Information Processing, SCI 182, pp. 95 – 118.
springerlink.com

computers to fill the semantic gap. In Information Processing, this paradigm shift has a great impact: users could provide input information in a perceptual form (e.g. in natural language) and could read and understand the subsequent results even without specific technical skills. In a nutshell, Human Centric Information Processing (HCIP) accounts users[1] as initiators of information processing as well as final recipients of the subsequent results. Through HCIP, machine intelligence increases dramatically and enables a more pervasive diffusion of computing.

The semantic gap between users and machines is due to the different nature of information that is represented and manipulated by these two actors. Two important features distinguish perceptual information from precise information, as pointed out by Zadeh: granularity and fuzziness (or graduality) [4-6].

Granularity refers to the property of information to refer to a clump of objects instead of a single one. Objects in a granular information (or information granule) are related by some proximity relation (in a wide sense). Representation and manipulation of information granules fall within Granular Computing, a key computing paradigm for HCIP [7,8].

Information granularity is required for economizing the representation of complex situations and phenomena, where precision is not necessary. Thus, granular information is used in mental processing of perceptual information. Furthermore, information granularity enables the use of natural language to describe facts. Most natural language sentences indeed represent granular information (e.g. “there is warm temperature in the room” does not specify any precise degree). This form of information could be sufficient for users to make decisions (e.g. turn-on the air conditioner), since in most cases users are unable to get more precise information (e.g. the exact temperature distribution of the room) nor they are interested.

Fuzziness is strictly related to information granules. According to this property, the membership of an object to an information granule is gradual rather than dichotomic. Fuzziness reflects the fact that natural phenomena are continue rather than discrete, and they are perceived by people with continuity. It is hence natural to assume that mental percepts reflect the graduality of the perceived phenomena. As a consequence, the semantics of natural language terms, which are used to symbolically describe perceptual information, embodies the fuzziness property.

Fuzzy information granules define information with granularity and fuzziness properties. They capture the key features of perceptual information and are naturally represented in natural language. Hence they constitute the basic underpinning for HCIP.

Fuzzy information granules should be also represented in computers and some mathematical machinery should be available in order to process this type of information. Fuzzy Set Theory (FST) provides such a machinery. Fuzzy Information Granules are represented as fuzzy sets, and fuzzy set operators are used to elaborate information.

Through FST, information processing can take form. Users can input perceptual information in natural language or in similar forms to a computer program. Such perceptual information is converted into fuzzy sets (a process called “m-precisiation” [9,10], where ‘m’ stands for ‘machine’). Fuzzy sets are processed according to program objectives, and results are usually represented as fuzzy sets. Finally, resulting

[1] Thorough this Chapter users are intended as human users.

fuzzy sets are converted into linguistic forms, according to a "m-imprecisiation" mechanism. Users are unaware that computers use numeric/symbolic operations, and computers are unaware that what they are elaborating are actually representations of perceptual information. Fuzzy information granules constitute a communication interface between two very different worlds (Fig. 1).

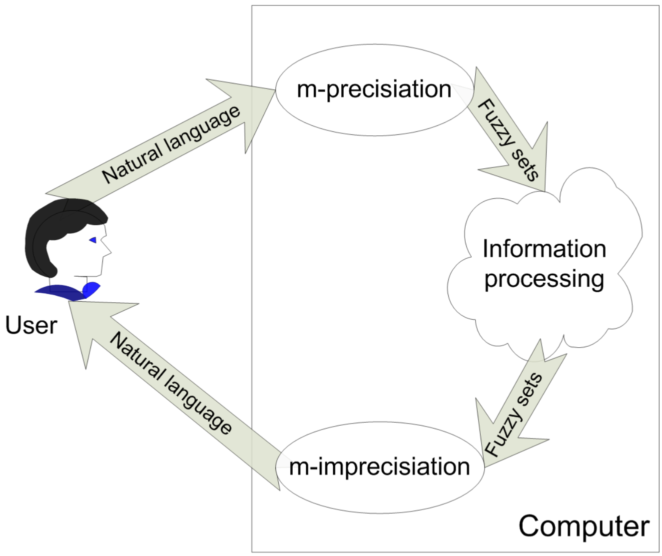


Fig. 1. Information flow in HCIP

We should note that without HCIP the same flow of information processing takes place, but with some remarkable differences: precisiation and imprecisiation are carried out by users (they are denoted with "h-precisiation" and "h-imprecisiation" in the Zadeh's notation, where 'h' stands for 'human') and information processing is carried out according to classical numerical/symbolic techniques.

Human-Centered information processing is not immune to problems directly deriving from the objects of processing, i.e. fuzzy information granules. More precisely, the processes of m-precisiation and m-imprecisiation are delicate tasks. The transformation of natural language terms into fuzzy sets and vice versa should be indeed "semantically nondestructive", i.e. the intrinsic semantics of a linguistic term and the explicit semantics of the corresponding fuzzy set should be highly overlapping[2]. This relation between semantics is called "co-intension" and tightly constrains the precisiation processes [11].

Despite their symmetry in scope, m-precisiation and m-imprecisiation are asymmetrical in the processes carried out. Usually, the process of m-precisiation is not as difficult as the process of m-imprecisiation. For m-precisiation, a number of reference fuzzy sets are usually available, and linguistic terms are converted into fuzzy sets with

[2] We cannot guarantee identity in principle because the semantics of a linguistic term is subjective and represented within the synapses of the brain in a imperscrutable way.

a direct matching. Such reference fuzzy sets are manually defined or automatically designed through information granulation techniques.

On the other hand, m-imprecisiation is more delicate. The reason is immediate: linguistic terms are usually limited, whilst the class of fuzzy sets is much wider. If the information processing task does not take into account co-intension in its computation, the resulting fuzzy sets cannot be easily associated to any linguistic term. In this case m-imprecisiation requires further elaboration so as to extract a convenient natural language description of the results with eventual information loss.

Fuzzy information granules that can be associated to linguistic terms are called *interpretable*. Interpretability of fuzzy information granules requires a deep understanding of the semantical constraints involved in the user-computer interaction. Issues regarding the role of interpretability, its definition, evaluation and preservation need to be addressed. The rest of this Chapter is devoted in eliciting some of those interpretability issues which may come up when designing a HCIP system. The next sections give a formal definition of fuzzy information granule as well as an attempt to define interpretability in a very general sense. Then, a number of methodological issues are discussed regarding modeling with interpretable information granules. Interpretability constraints are then discussed and finally an outline of interpretable information granulation strategies is reported. The chapter ends with some concluding remarks highlighting open issues and future trends on this topic.

2 Fuzzy Information Granules

An information granule is a “clump of objects which are drawn together by indistinguishability, similarity, proximity or functionality” [9]. An information granule arises from a process of “granulation”, which refers to the mental act of dividing a whole into semantically significant parts.

The definition of information granule is open to several formalizations, e.g. intervals, rough sets, fuzzy sets, etc. [7]. In particular, a fuzzy information granule is defined through fuzzy sets on the domain of considered objects. Fuzzy Set Theory is hence the mathematical underpinning of the Theory of Fuzzy Information Granulation (TFIG). The difference in the two theories is primarily epistemic: FST is a pure mathematical theory, TFIG has more semantic concerns since the relation that keeps together objects in the same granule is fundamental for the definition of an information granule.

For complex domains (i.e. domains of objects characterized by several attributes), fuzzy information granules can be employed to represent pieces of knowledge. By assigning a name to each attribute, a fuzzy information granule can express a soft relationship between two or more attributes. A collection of fuzzy information granules can be used as a knowledge base to make inferences about the possible values of an attribute when the values of other attributes are given.

More formally, suppose that the object domain is $U = X \times Y$ (sub-domains X and Y could be multi-dimensional as well, but here we are not interested in the nature of such sets). A fuzzy information granule is completely represented by a fuzzy set on U, i.e. a fuzzy relation between objects in X and Y. We denote with Γ such a fuzzy set; we write $\Gamma : U \to [0,1]$ or, equivalently, $\Gamma : X \times Y \to [0,1]$. A knowledge base (also called "granular world" [7]) formed by fuzzy information granules $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ is defined by accumulation of the pieces of knowledge defined by each single fuzzy information granule. This is a direct consequence of the definition of granulation, intended as a division of the whole (the domain *U*) into parts (the granules Γ_i).

The knowledge base (KB) is hence defined by the union of all the constituting fuzzy information granules, i.e.:

$$\mathrm{KB} = \bigcup_{i=1}^{n} \Gamma_i \tag{1}$$

To make inference, an information granule is interpreted as a possibility distribution over two variables *x* and *y*. When *x* is assigned a value (i.e. $x = \bar{x}$), inference is carried out by computing the possibility distribution of variable *y* for each information granule, namely π_i^y, such that:

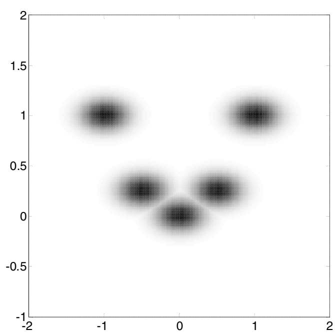
$$\pi_i^y(y) = \Gamma_i(\bar{x}, y) \tag{2}$$

A very common case is when each granule Γ_i is defined as the Cartesian Product of two fuzzy sets, namely $A_i : X \to [0,1]$ and $B_i : Y \to [0,1]$ and $\Gamma_i(x, y) = A_i(x) * B_i(y)$ being '*' a t-norm. In this case $\pi_i^y(y) = \Gamma_i(\bar{x}, y) = A_i(\bar{x}) * B_i(y)$. When the possibility distributions of all fuzzy information granules have been computed, the final possibility distribution for the entire KB is defined as $\pi^y(y) = \max_{i=1}^{n} \pi_i^y(y)$.

A KB made of fuzzy information granules is also called fuzzy graph [9]. Such structure is commonly used for approximating partially known functions, and is usually expressed in terms of a set of fuzzy rules, i.e. formal structures expressed as:

IF x is L[A] THEN y is L[B]

being L[A] and L[B] formal labels for fuzzy sets A and B respectively, and with the convention that "THEN" does not mean implication (as in logical rules) but it is a



If x is about -1 then y is about 1
If x is about -½ then y is about ¼
If x is about zero then y is about 0
If x is about ½ then y is about ¼
If x is about 1 then y is about 1

(a) (b)

Fig. 2. A granular parabola defined by five information granules: (a) graphical representation; (b) fuzzy rule representation

conjunction (it should be read "and then"). In fig. 2 a fuzzy graph representing a granular parabola is depicted.

Different interpretations of fuzzy rules are actually possible [12]. However, the interpretation of fuzzy rules as pieces of a fuzzy graph seems the most natural in the context of TFIG because it adheres to the definition of granulation and because the inference process is a direct application of the Compositional Rule of Inference.

We should note that the operator '*' used for combining fuzzy sets into an information granule is usually a t-norm (the functional representation of the conjunction) and the aggregation of information granules into a knowledge base is achieved through a t-conorm (the functional representation of the disjunction). This choice is conformant with the definition of information granule and the process of granulation as a division of a whole into parts. Here the domain U is divided into information granules (Γ_i), hence the t-conorm acts as an aggregation operator that merges separate pieces of information.

On the other hand, the t-norm used for defining each $\Gamma_i = A * B$ is conformant with the definition of information granule, intended as a clump of objects kept together by a tying relationship. Indeed, according to the definition, an object (x,y) is in the granule Γ_i if x is A AND y is B. If the t-norm is replaced by a t-conorm, the granule Γ_i could be split into two granules $\Gamma_{1,i}$ and $\Gamma_{2,i}$ where an object (x,y) belongs to $\Gamma_{1,i}$ if x is A (and y is any), or (x,y) belongs to $\Gamma_{2,i}$ if y is B (and x is any). The need of a t-norm is a consequence of the use of a t-conorm for merging fuzzy information granules, which in turn is a consequence of the definition of the process of granulation as a division of a whole into parts.

With this in mind we should exclude logical rules (prolog like) as representing information granules, because these rules represent implication in the material form x is NOT A OR y is B. Actually information granules grasp the core of a rule, i.e. when both the antecedent and the consequent are valid. The availability of more information granules enables correct inference when the antecedent in an information granule is not verified. To validate inference, it is assumed that all the antecedents of rules cover the domain of input x.

Since each information granule grasps only the core of a rule, it is closer to the way human beings reason in terms of rules. It is well-known that material implication is sometimes counterintuitive, because of the validity of material implication when the antecedent is not verified (see, e.g. [13]). On the other hand, information granules represent pieces of knowledge that are more co-intensive with knowledge of users. In this sense, information granules seem a good form of knowledge representation for HCIP.

3 Interpretability of Fuzzy Information Granules

Information granules are the basic blocks for communicating information and knowledge from users to machines and vice versa. When communication starts from users and is directed to machines, information granules are built by a precisiation procedure, which is aimed at finding the most co-intensive representation of the perceptual information that is wished to be communicated to the machine. Several approaches have been proposed in literature, especially for fuzzy information granules. Here the key issue is the elicitation of the membership function of the fuzzy set representing the input (see, e.g. [14, Ch. 3]).

Different is the case when machines should communicate information granules. Such granules could result from deductive inference process, or could emerge from a process of inductive learning, e.g. after data clustering processes. In both cases the main problem is to give a representation of the information granule that could be easily understood by the recipient of the communication, i.e. it is co-intensive with some known concepts hold by the user.

When the information granule is a result of inference, usually defuzzification is applied, in order to reduce the granule to a single element of the domain (prototype). There is no doubt that this method is the most direct but the most wasteful. Information about the specificity (precision) of the information granule is lost, and usually no information is provided about the significance of the chosen prototype [15].

When information granules arise after a clustering process, the quest for interpretability becomes more stringent. Many clustering schemes exist in literature, which are able to find data clusters of several shapes. However, in the context of information granulation, the accuracy of the clustering process is only one of two facets: interpretability should be taken into account as well.

Interpretability is crucial in HCIP, especially when knowledge has to be extracted from data and represented in a comprehensible form. Interpretability is necessary to easily and reliably verify the acquired knowledge and to relate it to user's domain knowledge, to facilitate debugging and improving the granulation technique; to validate granules, for their maintenance, and for their evolution in view of changes in the external world [16-18]. This is especially important when the domain knowledge and the discovered knowledge must be merged together (e.g. in knowledge intensive systems) [19]. Finally, and maybe most importantly, interpretability is needed for convincing users that the model is reliable, especially when they are not concerned with the techniques underlying the granulation process. Users of a decision support system should be confident on how it arrives to its decisions. This is particularly important in domains such as medical diagnosis [20].

3.1 A Definition for Interpretability

Most interpretability-oriented modeling techniques adopt an interpretation of the "Occam's Razor" principle. The spirit of this approach is to guarantee interpretability by simplifying the description of the involved information granules. Several works on clustering go in this direction [21]. Whilst necessary, the accordance to the Occam's Razor principle misses the point of interpretability, i.e. co-intensivity with user knowledge, which is unquestionably a richer and more complex requirement than simplicity.

In order to get a deeper insight on interpretability, a suitable definition should be given to this quality. A suitable characterization for interpretability is given by the so-called "Comprehensibility[3] Postulate", proposed by Michalski, a scholar in the Machine Learning community [22]. The Comprehensibility Postulate states that:

[3] Throughout the chapter the terms "interpretability", "comprehensibility" and "understandability" are considered as synonyms.

> The results of computer induction should be symbolic descriptions of given entities, semantically and structurally similar to those a human expert might produce observing the same entities. Components of these descriptions should be comprehensible as single “chunks” of information, directly interpretable in natural language, and should relate quantitative and qualitative concepts in an integrated fashion.

The key point of the Comprehensibility Postulate is the human-centrality of the results of a computer induction process. According to the postulate the results of computer induction (e.g. information granulation) should be described symbolically. Symbols are necessary to communicate information and knowledge. Pure numerical methods, including neural networks, are hence not suited for meeting understandability unless an interpretability oriented post-processing of acquired knowledge is performed, such as in [16,23].

Symbolic descriptions are necessary but might not be sufficient. They should be structurally and semantically similar to those a human expert might produce observing the same entities. This means that highly complex mathematical relationships, though described symbolically, may not be interpretable because they may not be compatible with human cognition. In the same way, knowledge logically represented by a huge number of rules (or predicates, clauses, etc.) do not meet the understandability feature, since humans have a limited ability to store information in short-term memory [24]. This passage suggests simplicity as a necessary, albeit not sufficient, condition for interpretability.

According to the postulate, symbols (or structures of symbols) should represent chunks of information. Here we recognize information granules as the semantic counterparts of symbols used for communication. In order to be understandable, symbols should be directly interpretable in natural language. This does not necessarily mean that symbols should be chosen from a natural language vocabulary, but has more profound implications. In particular, the Comprehensibility Postulate requires the interpretation of symbols to be in natural language. This is a requirement on the semantics of the symbols, i.e. on the information granules they denote. Therefore, in order to be understandable, information granules should be conformed with concepts a user can conceive.

We further observe that natural language terms convey implicit semantics (which also depends on the context in which terms are used), that are shared among all human beings speaking that language. As a consequence, a symbol coming from natural language can be used to denote an information granule only if the implicit semantics of the symbol highly matches with the semantics characterized by the information granule.

Finally, as the Comprehensibility Postulate requires, the description of computer induction results should relate both qualitative and quantitative concepts in an integrated fashion. We recognize in this requirement the role of TFIG as the most suitable candidate for representing information [4]. Fuzzy information granules can indeed represent both quantitative information (e.g. through fuzzy numbers, fuzzy intervals, fuzzy vectors, etc.) and qualitative information (usually represented as adjectives such as “low”, “medium”, “high”, “etc”). Both types of information have a homogeneous representation and could be elaborated in an integrated fashion.

3.2 Interpretability and Natural Language

We should emphasize that the requirement of natural language descriptions becomes more apparent as the complexity of the knowledge to be communicated increases. This is in coherence with the well known Zadeh's incompatibility principle [25]. When the knowledge to be communicated is simple, natural language is not strictly necessary. For example, linear regression results or interval-based rules are easy to understand even though they are usually represented in a mathematical form. Actually, even these forms of representation are in coherence with the Comprehensibility Postulate.

For example a linear equation coming from regression is actually a prototype of a granule including all possible data distributions that can be approximated by the line, i.e. it represents a chunk of information. Furthermore, the simple structure of the linear equation can be directly described in natural language. For example, the following linear model:

$$y = 4.323432122x + 1.98726325$$

can be described as:

y is proportional to x with factor about 4. For x=0 the value of y is about 2

As another example, consider an interval-based rule, such as:

$$\text{IF } x \in [2.372138, 4.675121] \text{ THEN } y \in [0.061115, 1.512143]$$

Again, when trying to understand this rule, a user may not focus on its finest details. Rather, she would stop on a more abstract descriptive level that depends from the context. For example, the user would understand that:

Whenever the value of x is about between 2.3 and 4.7, the value of y becomes smaller than about 1.5

Even though natural language description is not necessary to communicate this forms of knowledge, high level concepts are actually formed in the user mind when trying to understand the results of computer induction processes. In all cases, what the user builds in her mind could be described in natural language. As the complexity of the model increases, any precise representation becomes less and less comprehensible. For high levels of complexity, natural language seems to be the only mean to communicate knowledge and to make it understandable by users.

3.3 Interpretability and Information Visualization

We note that often users understand problems if the information is properly visualized in some form. Indeed, visualization has been recognized as a viable mean to enable users to interpret large amounts of data and to gain deeper insight into the working of complex systems. Visualization has been extensively investigated to pursuit understanding of complex patterns or models. Recently, more attention has been devoted to develop approaches to visualize fuzzy data and fuzzy models. [26]. Some of these approaches have the primary objective of helping users in understanding how a model works to generate its behavior. Other visualization techniques are mainly aimed at graphically representing knowledge so that users could easily interpret them.

All such techniques may offer an invaluable help to users in understanding induction results, even if they may not involve the Comprehensibility Postulate as the final representation is not symbolic. We observe, however, that visualization techniques may not fulfill the understandability requirement of information granules. Indeed, they are very useful for understanding how a behavior is generated, but the user may not understand why such behavior is correct, in the sense of providing significant outcomes. Furthermore, the main reasons that justify the interpretability features may not be fulfilled by visualization tools. In this situation, visualization techniques are complementary to the Comprehensibility Postulate, rather than alternative.

4 Interpretability Issues

Interpretability of information granules is a complex requirement that needs a comprehensive analysis of all facets of the environment on which granules are developed and used. This analysis results in a number of issues to be addressed for fulfilling the interpretability requirement.

4.1 The Objective of Granular Model

A first issue to be addressed for interpretability is the objective of the granular model, which may have a twofold nature: descriptive and prescriptive.

When the scope of the model is to describe a phenomenon, a data structure, etc., a number of interpretability constraints should be adopted in order to meaningfully tag the information granules with linguistic labels (symbols). In many cases, however, the granules are also used to make inference for predictions concerning new data. Briefly speaking, information granules results become part of a prescriptive model in a decision support system.

In all situations where understandability is required, attention should be paid also on how predictions are inferred from information granules. Specifically, the inference process should be cognitively plausible, so as to convince users on the reliability of the derived decisions. This is a delicate step often left unaddressed.

As an example, let us consider a Mamdani Fuzzy Inference System (FIS) [27]. Mamdani rules are defined through fuzzy information granules, by separating input variables from output variables. Even though the rules embodied in these FIS are built by responding to all the interpretability requirements, the inference carried out by the system may not convince the users about the reliability of the derived decision. Usually, the output of a Mamdani FIS is attained by applying a defuzzification procedure to the inferred fuzzy set such as Center-of-Gravity (CoG). However, CoG may not have any plausible explanation in some domains (e.g. medical diagnosis). This type of defuzzification would not be convincing for the user (e.g. physicians) about the reliability of the inferred output. On the other hand, more plausible forms of defuzzification would provide for more plausible inferences [28].

4.2 Representation Structure

To achieve interpretability, the structure of knowledge plays a fundamental role. Roughly speaking, two classes of structures can be distinguished: single and multiple representation structures [29]. Single representation structures are based on a “flat” representation of the knowledge base, usually in a rule-based form. A different approach provides for a multiple representation of knowledge, where one representation (not necessary interpretable) is used to generate an accurate prescriptive model, while the other is used to describe knowledge in an interpretable form [30,31] Such dual representation has evidence in brain organization in which different areas are devoted to perception, action performing and natural language communication.

Information granules may offer valuable help in defining both single and multiple representation structures. Single representation structures are naturally derived by defining a rule for each information granule (as in fuzzy graphs), or by tying two information granules with an implicative connector. However, multiple representations are also possible by defining two or more levels of granulation of the same data. In this case, the top levels can be used for descriptive pursuits, while more accurate predictions could be taken through bottom level information granules, where interpretability is less stringent and hence information granules could take shapes more conformant to the underlying data structures. Techniques for context-based information granulation achieve multiple representation structures [33 Ch. 4, 34].

4.3 User Characterization

Interpretability concerns the characterization of the user accessing the knowledge base of a model. This characterization drives the choice of the most appropriate representation of information granules and, hence, the constraints required for granting their interpretability.

Users are mainly characterized by their needs. Some users might be interested in understanding the information granules derived from data in order to use them for their purposes. Other types of users could be more interested in the validity of information granules, especially in terms of predictive accuracy. For the second type of users, interpretability requirement is not as stringent as for users of the first type.

The verification of interpretability of information granules tightly constrains their shape. As a consequence, a strong bias is introduced in interpretable information granules, resulting in a weaker predictive accuracy w.r.t. information granules without interpretability requirements. The interpretability/accuracy tradeoff should be taken into account when designing a new system. The choice strongly depends on the need of the final users.

If users require understandability of information granules, then some guidelines should be followed. The domain expertise of the user helps in the choice of the most appropriate representation of information granules. For example, highly qualitative forms of representations might be useful for users who are primarily interested in understanding granulation results. On the other hand, more precise forms of representation would be more useful for users who make decisions on the basis of granulation results. Furthermore, the user vocabulary is helpful in the choice of the linguistic terms to be used for representing information granules. Finally, the required level of precision is important to choose the granularity level of information granules, as well as to decide for single or multiple levels of representation.

5 Interpretability Constraints

An important question concerns how to verify if an information granule is interpretable. The question is ill-posed because the definition of interpretability is blurry, subjective and context-dependent. However, a general approach can be set, which is mainly based on constraints.

Several interpretability constraints have been proposed in literature: a recent survey can be found in [35]. Some of them have a precise mathematical characterization, while others have a more fuzzy definition. This is expectable, since imprecise definitions of interpretability constraints may be more co-intensive with our perception of interpretability.

The choice of interpretability constraints mainly depends on user characterization, granulation objectives, etc. As already noted, any constraint imposed on information granules introduces new bias on their ability of representing data relationships. As a consequence, the choice of interpretability constraints should be as careful as possible.

Interpretability constraints can be organized in a hierarchy reflecting the level to which they are applied. A convenient hierarchy for fuzzy information granules is the following:

1. Constraints on one-dimensional fuzzy sets;
2. Constraints on frames of cognition[4];
3. Constraints on information granules;

In the following, a brief discussion on such constraints is reported. For a deeper argumentation, the reader is referred to [35].

5.1 Constraints on One-Dimensional Fuzzy Sets

In defining fuzzy information granules, we highlight their constitution as Cartesian products of one-dimensional fuzzy sets. This assumption helps to decompose information granules as conjunction of simpler properties, all characterized by one-dimensional fuzzy sets. Such fuzzy sets should be co-intensive with elementary concepts, usually represented in the form “v is A” being “v” the name of an attribute and “A” the name of a quality, whose semantics is defined by a fuzzy set.

In order to be co-intensive with elementary concepts, one-dimensional fuzzy sets should verify a number of constraints. The choice of such constraints is mostly driven by common-sense, as well as to avoid some paradoxical situations that can occur when they are violated.

5.1.1 Normality

All one-dimensional fuzzy sets should be normal, i.e. there exist atleast one element of the domain with full membership. Normality is required to avoid paradoxes such as

[4] A frame of cognition is intended as the set of all one-dimensional fuzzy sets defined for the same attribute.

inclusion in the empty set, as proved in [36]. However, this constraint is required for co-intensiveness since it is expected that qualities are always fully met by some elements of the domain. For unlimited domains, asymptotic normality is also acceptable (e.g. the concept "X is distant" could be represented in this way).

5.1.2 Convexity

A one-dimensional fuzzy set is convex if it is defined on an ordered domain (e.g. the real line). Convexity implies that any midpoint between two extremes has membership degree greater or equal to the minimum membership of the two extremes. Convex fuzzy sets are very common (e.g. triangular, trapezoidal, Gaussian, etc.) and they are widely used because they express the semantics of similarity-based qualities, i.e. all qualities for which sentences like "x is more A than y" make sense. Non-convex fuzzy sets can be used too, as proposed in [37], but usually they represent compound qualities, which should be denoted by complex linguistic labels (e.g. "mealtime", meaning "breakfast or lunch or dinner").

Strict convexity requires that the membership of any midpoint between two extreme points is strictly greater than the minimum membership of the two extremes. Gaussian fuzzy sets are strictly convex, while trapezoidal and triangular fuzzy sets are convex, but not strictly convex. Strictly convex fuzzy sets are suited for modeling concepts characterized by a magnitude. As an example, the concept "hot" is characterized by the perceived temperature. The degree to which a temperature is "hot" monotonically increases as the temperature increases. While a Gaussian fuzzy set effectively models such a relationship, a trapezoidal fuzzy set fixes a threshold after which any temperature is considered "hot" with the same degree. While justifiable for efficiency pursuits, the trapezoidal fuzzy sets may not as co-intensive as Gaussian fuzzy sets with the mental concept of "hot", whose degree of evidence presumably varies with continuity.

5.1.3 Continuity

A fuzzy set defined on the real line should be continuous, so as to avoid abrupt changes of membership for very close elements. Continuity is especially required when fuzzy sets model perceptual information derived by observing macroscopic physical phenomena. Crisp set are examples of non-continuous fuzzy sets, leading to well known boundary paradoxes that make them unsuitable for modeling perceptual information and knowledge.

5.2 Constraints on Frames of Cognition

In any modeling context the number of fuzzy sets considered for each attribute is quite limited. The collection of all used fuzzy sets for a given attribute is named Frame of Cognition (FOC), as proposed in [36]. Linguistic Variables (defined in [38]) include a FOC if the number of generable linguistic values is finite. When a FOC is considered, a number of constraints should be verified in order to guarantee interpretability.

5.2.1 Proper Ordering

Linguistic terms used in natural language are related each other by two main relations: order and inclusion. As an example, the ordering of linguistic terms "small", "medium" and "high" is quite intuitive. Also, in some contexts we understand that the semantics of "very L" is included in the semantics of the linguistic term L. As remarked in [39], interpretable fuzzy sets within a FOC should reflect such basic relationships.

Inclusion of fuzzy sets is straightforward: A is included in B iff the membership degree of any element to A is less than to B. Fuzzy sets defined on an ordered domain could be also partially ordered in the following way: given two fuzzy sets A and B, we say A<=B iff there exists a midpoint *t* such that each point less than *t* has membership to A greater than to B, and each point greater than *t* has membership to B greater than to A. Thus A better represents elements of the Universe of Discourse that are smaller than the elements represented by B. In this sense, the ordering of fuzzy sets reflects the semantics formalized by their membership functions. If this constraint is violated, undesired situations may occur, which hamper interpretability. As an example, given a FOC with two fuzzy sets "cold" and "hot", we expect that for high temperatures the membership to hot is greater than the membership to cold, and vice versa for lower temperatures. This constraint calls for a proper choice of fuzzy sets in a FOC (see fig. 3 as illustration of an incorrect choice of fuzzy sets).

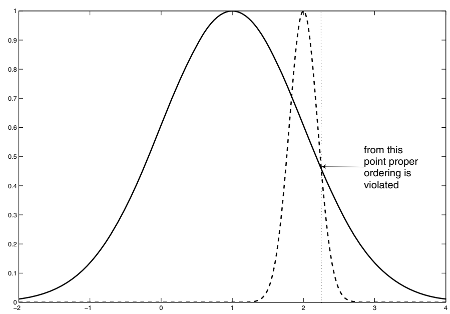


Fig. 3. Example of two Gaussian fuzzy sets violating proper ordering

5.2.2 Justifiable Number of Elements

In designing interpretable FOC, the number of fuzzy sets should be kept as small as possible, so that users could easily give appropriate meanings to the linguistic terms. By limiting the number of fuzzy sets in a FOC, a user is able to remember the proposed partition of the attribute domain. This greatly enhances interpretability.

The number of fuzzy sets is usually limited to 7±2, according to some psychological experiments reported in [24] showing the limited capacity of our short term memory in storing information. This limit has been debated (see [67] for a comprehensive discussion), and even smaller limits have been found in more recent experiments on immediate memory. Yet, a definitive answer on the capacity of primary memory has not been given, but psychological experiments (and common sense) suggest to keep the number of elements to remember very small.

The criterion of justifiable number of element spans all objects in a granular model. It is applicable to fuzzy sets in a FOC, as well as to the fuzzy sets compounding an information granule and to the number of information granules within a model. This extension is possible because our short-term memory is able to store simple structures (such as the names of fuzzy sets) as well complex structures (such as granule descriptions), provided that they are in small number.

We note that this criterion provides for a sound explanation of all simplification routines that are usually applied after clustering processes to enhance interpretability (see, e.g. [40] for a recent approach). However, we note also that this constraint severely limits the degrees of freedom (i.e. the free parameters) of a model. As a result, an interpretable model is highly biased and the resulting accuracy could be worse than an interpretability-free model. For this reason, interpretability is a feature that should be included with care in the design process.

5.2.3 Distinguishability

Roughly speaking, distinguishable fuzzy sets are well disjoint so they represent distinct concepts and can be assigned to semantically different linguistic labels. Well distinguishable fuzzy sets are deemed important since they obviate the subjective establishment of membership-function/linguistic term association, as claimed in [41], and reduce potential inconsistencies and redundancies in fuzzy models, as shown in [42]. Most importantly for the interpretability side, distinguishable fuzzy sets ease the linguistic interpretation of the model since fuzzy sets represent well separated concepts.

Distinguishability is a relation between fuzzy sets that can be formalized in several ways. Usually, a similarity measure between fuzzy sets is used, but the possibility measure can be also used under certain conditions, as showed in [43]. Possibility measure usually depends on the parameters of the membership functions, hence its calculation might be more efficient than similarity.

5.2.4 Coverage

The coverage constraint requires that every element of the domain belongs to at least one fuzzy set. Since membership is a matter of degree, the coverage constraint could be weak (membership greater than zero) or strong (membership greater than a threshold). In the latter case, the term α-coverage is used, being α a threshold in $]0,1[$.

Coverage is related to completeness, a property of deductive systems that has been used in the context of Artificial Intelligence to indicate that the knowledge representation scheme can represent every entity within the intended domain [44]. In [45] coverage (there called "cover full range") is justified by the fact that in human reasoning there will never be a gap of description within the range of the variable. On the contrary, as shown in [46] incompleteness may be a consequence of model adaption from data and can be considered a symptom of overfitting.

For the pursuits of interpretability, 0.5-coverage is desirable. This threshold corresponds to the optimal α-cut when fuzzy sets are converted into crisp sets, as proved in [36]. This means that if 0.5-coverage is not guaranteed, then for some elements of the domain the FOC represent only negative qualities (e.g. x is not hot and not cold, see fig. 4). Usually, natural language has positive terms to describe such elements (e.g x is warm). As a consequence the inclusion of fuzzy sets in the FOC – so that 0.5-coverage is guaranteed – enhances the interpretability of the granular model.

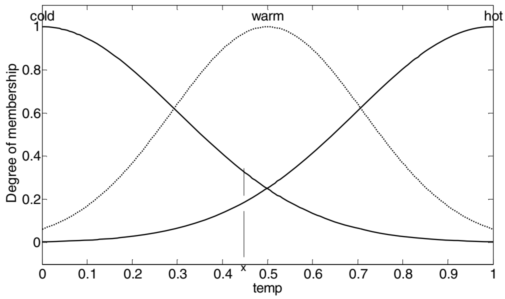


Fig. 4. Coverage and interpretability. Without fuzzy set labeled "warm", values around 0.5 of the variable domain do not have a linguistic description but they can only be described as "not cold and not hot".

5.2.5 Representation of Special Elements

For the pursuits of interpretability, it is often required that special elements of the universe of discourse are prototypes of some fuzzy sets in the FOC. In this way, such special elements are fully covered by some fuzzy sets which, in turn, represent special concepts.

Examples of special elements are extreme points of the universe of discourse. Leftmost and rightmost elements should be prototypes of some fuzzy sets that could be labeled in order to express their limit position in the FOC (e.g. "low", "high", "left", "right", etc.). The definition of such fuzzy sets is important to avoid paradoxical situations such as those depicted in fig. 5a. If the three fuzzy sets represent concepts "small", "medium", "high" then the leftmost element minU is less "small" than minU+e. Such an undesired situation will not occur if leftmost/rightmost fuzzy sets are defined in the FOC (fig. 5b). Leftmost and rightmost fuzzy sets are necessary when they represent qualitative concepts; if the FOC is made of fuzzy sets expressing fuzzy quantities, then they are not necessary.

Other special elements could be considered as prototypes of some fuzzy sets. The choice of such elements is problem driven. As an example, in [44] it is suggested that for control applications the null value, if it belongs to the universe of discourse,

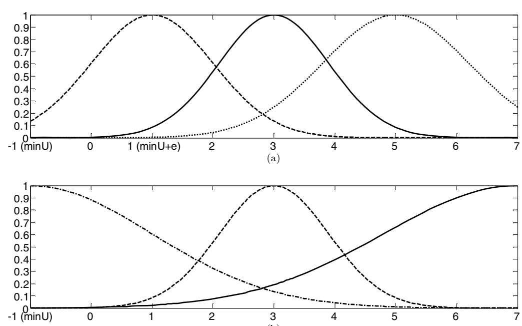


Fig. 5. (a) A FOC violating representativity of extreme values. (b) a FOC with leftmost/rightmost fuzzy sets.

should be prototype of a fuzzy set, expressing the concepts “nearly zero”. This suggestion could be extended to other values deemed important in an applicative context. For example, 0 and 100 could be prototypes of fuzzy sets labeled “icing point” and “boiling point” in the domain of water temperatures (expressed in Celsius degrees), or 37°C could be considered as prototype in a FOC expressing human body temperatures.

5.3 Interpretability Constraints on Information Granules

A fuzzy information granule is defined as a Cartesian product of fuzzy sets, each coming from a different FOC. The relational nature of information granules make them the basic building blocks for expressing knowledge. To make such knowledge comprehensible, a number of interpretability constraints should be verified.

5.3.1 Justifiable Number of Elements

The constraint of justifiable number of elements (JNE in brief) is applied at all levels of granulation: the FOC, each single information granules and the collection of information granules in a model.

When considering a single information granule, the JNE constraint imposes that the number of fuzzy sets defining the granule should be kept small, e.g. less than about seven. In this way, it is easier for the user to build a mental concept associated to the granule. Very complex granules are difficult to understand and, even if the compounding fuzzy sets verify all suggested interpretability constraints, users may not be able to grasp the relationship among variables that is represented by the granule. Techniques such as variable selection, locally to the granule or globally to the entire knowledge base, are useful to improve interpretability.

When considering the overall knowledge base, the JNE constraint suggests that the number of information granules should be kept within a small limit. According to psychological experiments, indeed, our short term memory is able to store simple as well as complex structures, provided that they are in a small number. Again, this poses severe limits on the flexibility of the resulting model, which may negatively influence its accuracy.

5.3.2 Completeness

This constraint imposes that each element of the universe of discourse is covered by at least one information granule, i.e. it belongs to an information granule with a membership degree greater than zero (weak completeness) or greater than a specified threshold (strong completeness).

Weak completeness is easy to achieve if fuzzy sets with infinite support are used (such as Gaussian fuzzy sets). On the other hand, strong completeness may pose some problems, especially when the specified threshold is high (e.g. 0.5 as usually required) and the dimensionality of the universe of discourse is high. In this case, indeed, a great number of information granules may be required, which may hamper the constraint of justifiable number of elements. To overcome this problem a form of “closed world assumption” is usually adopted. It is assumed that all elements that can occur are represented by at least one information granule with membership degree greater than a threshold. According to this assumption, the universe of discourse is actually a subset of the Cartesian product of all attribute domains, and it is assumed that this subset is covered by the information granules.

A drawback of the closed world assumption emerges when an element not belonging to the subset occurs. In this case the model may infer weak results (i.e. highly subnormal fuzzy sets). To overcome this problem, in [47] a "default" information granule is used, whose membership function is defined as the complement of the union of all used information granules. In this way, any element not represented by any information granule is covered by this default granule, to which a special action could be attached.

5.3.3 Correctness

The correctness constraint applies to the inference process carried out by the granular model. Informally speaking, correctness imposes that the inference process provides logically consistent outputs. As an example, in rule-based models, correctness requires that Modus Ponens is respected, i.e. if a rule of the type "IF x is A THEN y is B" belongs to the rule base, and the input A is provided, it is expected that the model output is B.

On the basis of this definition, several efforts have been made to verify the correctness of rule-based models, such as in [48]. For granular models, we should keep in mind that the knowledge base is made by the union of information granules, and each information granule is defined by the conjunction of elementary concepts. As a consequence, if a granule representing "x is A and y is B" belongs to the model, and the input A is provided, then the model is expected to derive B possibly united with other fuzzy sets.

As an example, consider a model with two information granules, labeled as "temperature is very cold and position is north pole", and "temperature is very cold and position is south pole". If the fact "temperature is very cold" is provided, we should expect the inference "position is north pole OR south pole". In this sense, rule inconsistencies are not possible in granular models (see also [49] for a formal treatment of the topic).

6 Interpretable Fuzzy Information Granulation

Information granulation is the process of discovering granules from data by extracting hidden relationships among observed samples. The nature of such relationships depends on the granulation algorithm and defines the semantics of the resulting information granules.

According to Zadeh [4], granulation is a cognitive task devoted to the partition of a whole into (significant) parts. Conversely, the process of aggregating parts into a whole is referred as organization. By virtue of such a definition, we may interpret granulation as the discovery of relationships of data *within* parts, so that the latter are semantically significant. On the other hand, organization involves the discovery of relationships *between* parts.

Interpretable information granulation adds interpretability constraints to the granulation process. The choice of which constraints to include is a matter of design. In the following, a number of commonly adopted strategies for interpretable information granulation is outlined. For a deeper review of interpretable granulation techniques (in the context of fuzzy modeling) the reader is referred to [50].

6.1 Partitioning

A widely adopted strategy for interpretable information granulation concerns the partition of the data domain into fuzzy granules that verify a number of interpretability constraints.

Partitioning can be fixed or dynamic. In fixed partitioning, a number of interpretable fuzzy sets is defined for each attribute (a FOC), and information granules are obtained by combining fuzzy sets of different attributes. To avoid combinatorial explosion of information granules, only those including an adequate number of available samples is retained, while all the others are discarded.

Fixed partitioning provides for very interpretable fuzzy information granules but suffers of many drawbacks. The main shortcoming derives from the definition of fuzzy sets, which does not take into account the structure of data. As a consequence fixed partitioning may not represent the most adequate granulation of data. Furthermore, the choice of the number of fuzzy sets for each attribute determines the granularity level of each information granule. Without any information of data distribution, an arbitrary choice of the level of granularity may seriously hamper the quality of the granulation process. Despite these drawbacks, however, manual partitioning is still widely used because of its simplicity [63].

To avoid the shortcomings of fixed partitioning, dynamic partitioning techniques have been proposed. The fundamental strategy of dynamic partitioning is to refine an initial partition so as to better represent data relationships, without violating interpretability constraints. Refinement usually applies merge and split operators for fuzzy sets [51, 64], or modification of fuzzy set parameters [52], or both [53]. In [65] fuzzy sets in a FOC are defined by a frequentist approach so that more specific fuzzy sets are defined to cover attribute values with higher frequency.

Alternative to these partition strategies, some works use fuzzy tree-based partitioning to granulate data [54]. Roughly speaking, for generating a tree-based partition an attribute is selected and split in two fuzzy sets. For each of the two parts the split algorithm is applied on the remaining attributes. This approach leads to a compact representation of information granules (especially because attribute selection is usually performed), but the resulting granules may not share fuzzy sets. This implies a number of similar fuzzy sets to be defined for the same attribute, which may hamper interpretability (fig. 6).

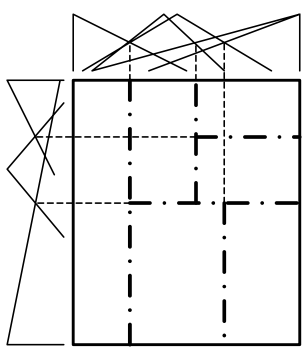


Fig. 6. Example of tree-based partition

6.2 Clustering

Clustering techniques are widely used for data granulation. This is due to the ability of clustering techniques to discover hidden relationships from data. Several fuzzy clustering techniques have been proposed in literature (see [55] for a review), however few of them address interpretability.

The main difficulty for assuring interpretability of fuzzy granules resulting from clustering processes generally stands in the difficulty of representing fuzzy clusters in natural language. A common approach to assure natural language representation is to express fuzzy clusters as Cartesian product of fuzzy sets. This however implies that the shape of fuzzy clusters is tightly constrained. Furthermore, the presence of several clusters may lead to a high number of very overlapped fuzzy sets for each attributes. This situation does not lead to an interpretable granulation of data.

To improve interpretability of cluster-based information granules, often simplification procedures are proposed, which merge similar fuzzy sets of the same FOC into single fuzzy sets [56]. This approach yields compact granular models, but often other interpretability constraints (e.g. coverage, representativity, etc.) are not fulfilled. They are hence most suited for quantitative information granulation, where granules represent imprecise quantities for each attribute.

When qualitative information granulation is required, i.e. when granules represent qualities on each attributes, a greater number of interpretability constraints should be verified. In this case, other strategies are advisable. In [57] an approach for interpretable granulation is proposed, which is based on a double clustering process. The first clustering stage operates on the entire dataset in order to discover hidden relationships among data. The result of this stage is a collection of prototypes that synthetically describe the dataset. In the second stage, multidimensional prototypes are projected onto each single attribute and further clustered to achieve the desired granulation level. One-dimensional projections are used to define FOCs that verify a number of interpretability constraints so that the resulting fuzzy sets can be labeled with qualitative linguistic terms. Such fuzzy sets are combined (one for each attribute) in order to define fuzzy information granules that represent data in a natural language form. Variants of the Double Clustering schema are also able to automatically determine the granularity level [58,59].

7 Concluding Remarks

The main objective of interpretability in fuzzy information granulation is co-intensiveness with human perceptual knowledge. Fuzzy information granules are basic building blocks for representing semantical knowledge in a computer-manageable form. Without interpretability, fuzzy information granules are still able to represent imprecise knowledge, similarly to the knowledge learned by a neural network or some other black box model, but its comprehensibility by users is limited, especially when they do not possess skills in fuzzy and granular technologies.

In this sense, interpretability is a fundamental feature for using fuzzy information granules in Human-Centric Information Processing. But interpretability is not a mathematical property, it is rather an epistemic feature that spans several facets of model design. Interpretability constraints are the formal counterparts of the interpretability property, provided that several other issues have been addressed, such as user characterization, representation structure and, last but not least, the objective of the model to be designed.

Current literature offers a great number of interpretability constraints, some of which have been revised in this chapter. Many of these constraints have been proposed as formalizations of properties driven by common sense. This opens the door to further research aimed at finding relationships (e.g. the "representation of special elements", which is a generalization of two or more constraints found in literature), at devising different formalizations of the same constraint (such as the "distinguishability" constraint or the "proper ordering") or at discarding some proposed constraint (such as the 1-complementarity of membership degrees, which has a technical rationale but cannot be justified in terms of interpretability). Even more importantly, groups of interpretability constraints help in identifying different notions of interpretability, such as interpretability of granules expressing quantities rather than qualities. These and other findings are important aids for knowledge engineering with fuzzy information granules.

Interpretability usually clashes with predictive accuracy. The more interpretability constraints are used, the more rigid is the granular model, and the less flexible to adaption it is. Interpretability vs. accuracy tradeoff has been addressed for long time, and several approaches have been proposed to balance these two features e.g. by regularized learning or multi-objective optimization (see, e.g. [54, 60]). Furthermore, the adoption of interpretability constraints should be carefully pondered in certain applications where accuracy has a prominent importance, such as in fuzzy control [61].

Current and future research on interpretability spans both methodological and theoretical issues. Among these, the representation of the semantics of natural language terms is of particular interest. Mendel [62] proposes type-2 fuzzy sets for such a representation. This is a promising research direction, which may result particularly fruitful in the area of granular knowledge communication (see also [66] for a discussion on this topic). On a more general level, we believe that deep insights on the semantics of membership degrees (which could denote similarity, preference, possibility of other, see [68]), as well as their operations, will shed new light on interpretability of information granules and, in turn, on Human-Centered Information Processing.

References

1. Pedrycz, W., Gomide, F.: Fuzzy Systems Engineering. Toward Human-Centric Computing. John Wiley & Sons, Hoboken (2007)
2. Pazzani, M.J.: Knowledge discovery from data? IEEE Intell. Syst. 15(2), 10–12 (2000)
3. Jaimes, A., Gatica-Perez, D., Sebe, N., Huang, T.S.: Human-Centered Computing–Toward a Human Revolution. IEEE Computer 40(5), 30–34 (2007)
4. Zadeh, L.A.: Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic. Fuzzy Sets and Syst. 90, 111–117 (1997)
5. Zadeh, L.A.: Graduation and granulation are keys to computation with information described in natural language. In: Proc. of IEEE GrC 2006, p. 30 (2006)
6. Zadeh, L.A.: Toward Human-Level Machine Intelligence. In: Proc. of SOFA 2007, pp. 9–10 (2007)

7. Bargiela, A., Pedrycz, W.: Granular Computing: An Introduction. Kluwer Academic Publishers, Dordrecht (2003)
8. Bargiela, A., Pedrycz, W.: Toward a Theory of Granular Computing for Human-Centered Information Processing. IEEE Trans. on Fuzzy Syst. 16(2), 320–330 (2008)
9. Zadeh, L.A.: Toward a generalized theory of uncertainty (GTU): an outline. Inf. Sci. 172, 1–2 (2005)
10. Zadeh, L.A.: Is there a need for fuzzy logic? Inf. Sci. 178(13), 2751–2779 (2008)
11. Zadeh, L.A.: A New Frontier in Computation – Computation with Information Described in Natural Language. In: Proc. of 2006 IEEE Conf. on Intell, pp. 5–6 (2006)
12. Dubois, D., Prade, H.: What are fuzzy rules and how to use them. Fuzzy Sets Syst 84(2), 169–185 (1996)
13. Brandom, R.: Semantic Paradox of Material Implication. Notre Dame J. of Form. Log. 22(2), 129–132 (1981)
14. Türksen, I.B.: An Ontological and Epistemological Perspective of Fuzzy Set Theory. Elsevier, Amsterdam (2005)
15. Roychowdhury, S.: An Inquiry into the Theory of Defuzzification. In: Pedrycz, W. (ed.) Granular Computing: an Emerging Paradigm, pp. 143–162. Physica-Verlag (2001)
16. Zhou, Z.H.: Comprehensibility of Data Mining Algorithms. In: Wang, J. (ed.) Encyclopedia of Data Warehousing and Mining, Hershey, PA, pp. 190–195 (2005)
17. Giboin, A.: ML Comprehensibility and KBS explanation: stating the problem collaboratively. In: Proc. of IJCAI 1995 Workshop on Machine Learning and Comprehensibility, pp. 1–11 (1995)
18. Mencar, C., Castellano, G., Fanelli, A.M.: On the role of interpretability in fuzzy data mining. Int. J. of Uncertain., Fuzziness and Knowl. Based Syst. 15, 521–537 (2007)
19. Michalski, R.S.: Knowledge Mining: A Proposed New Direction. In: Proc. of Symp. on Data Mining and Semantic Web (2003), `http://www.mli.gmu.edu/papers/2003-2004/03-5.pdf`
20. Kuncheva, L.I., Steimann, F.: Fuzzy diagnosis. Art. Intell. in Medicine 16(2), 121–128 (1999)
21. Guillaim, S.: Designing fuzzy inference systems from data: An interpretability-oriented review. IEEE Trans. Fuzzy Syst. 9(3), 426–443 (2001)
22. Michalski, R.S.: A theory and methodology of inductive learning. Artif. Intell. 20, 111–161 (1983)
23. Johansson, U., Niklasson, L., Knig, R.: Accuracy vs. Comprehensibility in Data Mining Models. In: Proc. of the 7th Int. Conf. on Inform. Fusion, pp. 295–300 (2004)
24. Miller, G.A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information. The Psychological Review 63, 81–97 (1956), `http://www.well.com/user/smalin/miller.html` (electronic reproduction by Stephen Malinowski)
25. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. on Syst., Man and Cybern. 3(1), 28–44 (1973)
26. Pham, B., Brown, R.: Visualisation of fuzzy systems: requirements, techniques and framework. Future Gener. Comput. Syst 21(7), 1199–1212 (2005)
27. Mamdani, E.H., Assilian, S.: An experiment in linguistic synthesis with a Fuzzy Logic controller. Int. J. of Human-Comput. Studies 51(2), 135–147 (1999)
28. Roychowdhury, R., Pedrycz, W.: A survey of defuzzification strategies. Int. J.of Intell. Syst. 16(6), 679–695 (2001)
29. van de Merckt, T., Decaestecker, C.: Multiple- Knowledge Representations in Concept Learning. In: Lavrač, N., Wrobel, S. (eds.) ECML 1995. LNCS (LNAI), vol. 912, pp. 200–217. Springer, Heidelberg (1995)

30. Domingos, P.: Knowledge Discovery Via Multiple Models. Intell. Data Analysis 2(3), 187–202 (1998)
31. Sun, R., Slusarz, P., Terry, C.: The Interaction of the Explicit and the Implicit in Skill Learning: A Dual-Process Approach. Psych. Rev. 112(1), 159–192
32. Stadler, M., Frensch, P.: Handbook of implicit learning. Sage, Thousand Oaks (1998)
33. Pedrycz, W.: Knowledge Based Clustering: from Data to Information Granules. John Wiley & Sons, Hoboken (2005)
34. Mencar, C., Castellano, G., Fanelli, A.M.: Balancing Interpretability and Accuracy by Multi-Level Fuzzy Information Granulation. In: Proc. of IEEE Int. Conf. on Fuzzy Syst., pp. 2157–2163 (2006)
35. Mencar, C., Fanelli, A.M.: Interpretability constraints for fuzzy information granulation. Inf. Sci. 178(24), 4585–4618 (2008)
36. Pedrycz, W., Gomide, F.: An Introduction to Fuzzy Sets: Analysis and Design. MIT Press, Cambridge (1998)
37. Garibaldi, J.M., Musikasuwan, S., Ozen, T., John, R.I.: A Case Study to Illustrate the Use of Non-Convex Membership Functions for Linguistic Terms. In: Proc. of IEEE Int. Conf. on Fuzzy Syst., vol. 3, pp. 1403–1408 (2004)
38. Zadeh, L.A.: The concept of Linguistic Variable and its application to approximate reasoning. Inf. Sci. 8, 199–249 (1975)
39. Bodenhofer, U., Bauer, P.: Interpretability of linguistic variables: a formal account. Kybernetika 41(2), 227–248 (2005)
40. Gegov, A.: Complexity Management in Fuzzy Systems: A Rule Base Compression Approach. Studies in Fuzziness and Soft Computing, vol. 211. Springer, Heidelberg (2007)
41. de Oliveira, J.V.: Semantic constraints for membership function optimization. IEEE Trans. on Syst., Man and Cybern. A 29(1), 128–138 (1999)
42. Roubos, H., Setnes, M.: Compact fuzzy models through complexity reduction and evolutionary optimization. In: Proc. of 9th IEEE Int. Conf. on Fuzzy Syst., vol. 2, pp. 762–767 (2000)
43. Mencar, C., Castellano, G., Fanelli, A.M.: Distinguishability quantification of fuzzy sets. Inform. Sci. 177, 30–149 (2007)
44. de Oliveira, J.V.: Towards neuro-linguistic modeling: Constraints for optimization of membership functions. Fuzzy Sets and Syst. 106, 357–380 (1999)
45. Hermann, C.S.: Symbolic reasoning about numerical data: A hybrid approach. Appl. Intell. 7, 339–354 (1997)
46. Jin, Y., Von Seelen, W., Sendhoff, B.: On generating FC3 fuzzy rule systems from data using evolution strategies. IEEE Trans. on Syst., Man and Cybern. B 29(6), 829–845 (1999)
47. Peña-Reyes, C., Sipper, M.: Fuzzy CoCo: Balancing Accuracy and Interpretability of Fuzzy Models by Means of Coevolution. In: Casillas, J., Cordon, O., Herrera, F., Magdalena, L. (eds.) Accuracy Improvements in Linguistic Fuzzy Modeling, pp. 119–146. Springer, Heidelberg (2003)
48. Perfilieva, I., Lehmke, S.: Correct models of fuzzy IF–THEN rules are continuous. Fuzzy Sets and Syst. 157(24), 3188–3197 (2006)
49. Dubois, D., Prade, H., Ughetto, L.: Checking the Coherence and Redundancy of Fuzzy Knowledge Bases. IEEE Trans. on Fuzzy Syst. 5(3), 398–417 (1997)
50. Guillaume, S.: Designing Fuzzy Inference Systems from Data: An Interpretability-Oriented Review. IEEE Trans. on Fuzzy Syst. 9(3), 426–443 (2001)
51. Guillaume, S., Charnomordic, B.: Generating an Interpretable Family of Fuzzy Partitions From Data. IEEE Trans. on Fuzzy Syst. 12(3), 324–335 (2004)

52. de Oliveira, J.V.: Towards neuro-linguistic modeling: Constraints for optimization of membership functions. Fuzzy Sets and Syst. 106(3), 357–380 (1999)
53. Nauck, D.D.: Fuzzy data analysis with NEFCLASS. Int. J. on Approx. Reason. 32(2-3), 103–130 (2003)
54. Mikut, R., Jäkel, J., Gröll, L.: Interpretability issues in data-based learning of fuzzy systems. Fuzzy Sets and Syst. 150(2), 179–197 (2005)
55. Baraldi, A., Blonda, P.: A survey of fuzzy clustering algorithms for pattern recognition. Parts I & II. IEEE Trans. on Syst., Man, and Cybern. B 29(6), 778–801 (1999)
56. Roubos, H., Setnes, M.: Compact and transparent fuzzy models and classifiers through iterative complexity reduction. IEEE Trans. on Fuzzy Syst. 9(4), 516–524 (2001)
57. Castellano, G., Fanelli, A.M., Mencar, C.: DCf: a double clustering framework for fuzzy information granulation. In: Proc. of IEEE Intern. Conf. on Granular Comput., vol. 2, pp. 397–400 (2005)
58. Castellano, G., Fanelli, A.M., Mencar, C., Plantamura, V.: Classifying data with interpretable fuzzy granulation. In: Proc. of SCIS&ISIS, pp. 872–877 (2006)
59. Mencar, C., Consiglio, A., Fanelli, A.M.: DCγ: Interpretable Granulation of Data through GA-based Double Clustering. In: Proc. of IEEE Int. Conf. on Fuzzy Syst., pp. 1–6 (2007)
60. Ishibuchi, H., Nojima, Y.: Analysis of interpretability-accuracy tradeoff of fuzzy systems by multiobjective fuzzy genetics-based machine learning. Int. J. of Approx. Reason. 44, 4–31 (2007)
61. Riid, A., Rustern, E.: Interpretability of Fuzzy Systems and Its Application to Process Control. In: Proc. of IEEE Int. Conf. on Fuzzy Syst., pp. 1–6 (2007)
62. Mendel, J.M.: Computing with words and its relationships with fuzzistics. Inform. Sci. 177, 988–1006 (2007)
63. Novák, V., Perfilieva, I., Dvořǎk, A., Chen, G., Wei, Q., Yan, P.: Mining pure linguistic associations from numerical data. Int. J. on Approx. Reason. 48, 4–22 (2008)
64. Herrera, F., Herrera-Viedma, E., Martínez, L.: A Fuzzy Linguistic Methodology to Deal With Unbalanced Linguistic Term Sets. IEEE Trans. on Fuzzy Syst. 16(2), 354–370 (2008)
65. Drobics, M., Himmelbauer, J.: Creating comprehensible regression models. Soft. Comput. 11, 421–438 (2007)
66. Dimitrov, V., Russell, D.: The Fuzziness of Communication. Fell L, Russell D, Stewart A, Seized by Agreement, Swamped by Understanding. Hawkesbury Printing, University of Western Sydney. (1994), `http://www.univie.ac.at/constructivism/pub/seized/fuzcom.html`
67. Cowan, N., Morey, C.C., Chen, Z.: The legend of the magical number seven. In: Della Sala, S. (ed.) Tall Tales about the Mind and Brain - Separating Fact from Fiction, pp. 45–59. Oxford University Press, Oxford (2007)
68. Dubois, D., Prade, H.: The three semantics of fuzzy sets. Fuzzy Sets and Sys 90(2), 141–150 (1997)