CrossMark

ORIGINAL PAPER

Granular computing, computational intelligence, and the analysis of non-geometric input spaces

Lorenzo Livi[1] · Alireza Sadeghian[1]

Abstract Data granulation emerged as an important paradigm in modeling and computing with uncertainty, exploiting information granules as the main mathematical constructs involved in the context of granular computing. In this paper, we comment on the importance of data granulation in computational intelligence methods. Toward this aim, we discuss also the peculiar aspects related to the analysis of non-geometric patterns, which have recently attracted considerable attention by researchers. As a conclusion, we elaborate over the fundamental, conceptual problems underlying the process of data granulation, which drive the quest for a sound theory of granular computing.

Keywords Data granulation · Computational intelligence · Analysis of non-geometric patterns · Uncertainty measures

1 Introduction

Granular computing (Pedrycz and Chen 2014; Pedrycz 2013; Han and Lin 2010; Bargiela and Pedrycz 2008) is typically portrayed as a research context intended as a convergence of various modeling and computational approaches for dealing with uncertainty. The modeling side is essentially rooted on formal constructs called information granules (IG) (Pedrycz et al. 2015, 2008). Information granules are mathematical models describing data aggregates; data in such aggregates are related to each other by considering, for instance, functional and structural similarity criteria. Well-known formal settings for implementing information granules include (higher-order) fuzzy sets (Wagner et al. 2015), rough sets (Qian et al. 2010; Ali et al. 2013), and intuitionistic sets (Huang et al. 2013), for instance. Information granules can be obtained in a data-driven fashion in different ways (Yao et al. 2013; Salehi et al. 2015). A prominent example comes from partition-based approaches, which are typically implemented by means of partitive clustering algorithms. However, formation of information granules is not limited to partition-based techniques (Salehi et al. 2015; Qian et al. 2014, 2015). Uncertainty is a pivotal concept in granular computing and related aspects. In fact, one of the main goals of information granules is to convey the uncertainty of aggregated data in a synthetic but yet effective way. Uncertainty is a powerful concept, which has been highly exploited in many mathematical settings (Klir 2006). Information theory is certainly the most prominent example, which is rooted in the classical probability framework. However, information-theoretic concepts have been extended to modern models of information granules, such as in the case of imprecise probabilities (Bronevich and Klir 2010), fuzzy sets (Zhai and Mendel 2011), intuitionistic fuzzy sets (Montes et al. 2015), and rough sets (Zhu and Wen 2012; Chen et al. 2014; Dai and Tian 2013).

The computational aspects of granular computing are intimately related with the research context called computational intelligence (Livi et al. 2015). It includes nature-inspired techniques for performing recognition, control, and optimization tasks (Engelbrecht 2007). Related methodologies are typically data-driven, in the sense that such techniques rely on experimental evidence (data,

✉ Lorenzo Livi
llivi@scs.ryerson.ca

Alireza Sadeghian
asadeghi@ryerson.ca

[1] Department of Computer Science, Ryerson University, 350 Victoria Street, Toronto, ON M5B 2K3, Canada

Springer

patterns) in order to perform inductive inference (or generalization). In this setting, information granules are used as computational components of data-driven inference systems. Granular neural networks (Ding et al. 2014) are only one of the many pertaining examples available in the literature (Pedrycz 2013).

In this paper, we comment on the importance of data granulation in Computational Intelligence methods. Section 2 introduces the Computational Intelligence context, posing the accent on pattern recognition aspects. We discuss also issues related to the analysis of the so-called non-geometric patterns, which have recently attracted considerable attention by researchers (Livi et al. 2015). Section 3 offers a glimpse of the rapidly-changing granular computing domain. We highlight two specific aspects in this paper: the (novel) interpretation of information granules as patterns (Sect. 3.1) and the challenge of designing a criterion for synthesizing information granules from data (Sect. 3.2). The design of such criteria is deeply connected with the fundamental, conceptual problems underlying the process of data granulation, which drive the quest for a sound theory of granular computing bridging both model-based and data-driven perspectives. Finally, Sect. 4 concludes the paper.

2 Computational intelligence methods and the challenge for processing non-geometric input spaces

The research context called computational intelligence (CI) (Engelbrecht 2007) unifies several nature-inspired computational methods under a data-driven paradigm. Well-known instances of such methods include neural networks, fuzzy systems, evolutionary and swarm intelligence optimization techniques. Typical problems faced by means of CI methods include recognition problems (e.g., classification, clustering, and function approximation) and those of adaptive control (e.g., fuzzy control and data-driven optimization via neural networks). CI is closely-related to the soft computing discipline. Quoting from Bonissone (Bonissone 1997) "The term soft computing (SC) represents the combination of emerging problem-solving technologies such as fuzzy logic (FL), probabilistic reasoning (PR), neural networks (NNs), and genetic algorithms (GAs). Each of these technologies provides us with complementary reasoning and searching methods to solve complex, real-world problems". Data-driven inductive inference systems can be implemented in terms of soft computing methodologies by departing from the assumption of Boolean truth values and membership of elements to classes. This goal was first obtained by means of the celebrated Zadeh's fuzzy logic (Zadeh 1965). Different many-valued logics (and corresponding set-theoretic frameworks) have been defined so far, such as rough sets (Pawlak 1982), intuitionistic fuzzy sets (Atanassov 1986), and the three-valued logic underlying shadowed sets (Pedrycz 1998). Well-known applications of fuzzy logic include rule-based (adaptive) fuzzy inference systems (Nauck et al. 1997), modern evolutions of fuzzy neural networks (Wu et al. 2014; Liu an Li 2004) and higher-order fuzzy systems (Zhou et al. 2009; Pagola et al. 2013; Wagner and Hagras 2010; Melin and Castillo 2013; Oh et al. 2014).

Focusing on recognition problems (Theodoridis and Koutroumbas 2008; Haykin 2007), the concept of pattern plays an important role. Patterns are everywhere, such as in climate physics, series of seismic events, complex biochemical and biophysical processes, brain science, financial markets and economical trends, large-scale power systems, and so on. Human knowledge and reasoning are both founded on searching for such patterns and on their effective aggregation to define meaningful concepts and decision rules (Pedrycz 2013). However, formally speaking, a pattern is essentially an experimental instance of a data generating process, P. A process can be described as a mapping $P : \mathcal{X} \rightarrow \mathcal{Y}$, idealizing a system (either abstract or physical) that generates outputs according to inputs. $\mathcal{X}$ is referred to as the input space (or domain, representation space), where the patterns are effectively represented according to some suitable formalism. $\mathcal{Y}$, instead, is the output space. In pattern recognition, the closed-form expression of P is not known. Nonetheless, it is possible to observe such a process through a finite dataset, $\mathcal{S}$. The problem typically boils down to reconstructing a mathematical model of P, say M, by analyzing $\mathcal{S}$. To be useful in practice, a mathematical model M, once established, must be adapted to the problem at hand. In practice, learning or synthesizing a model M from $\mathcal{S}$ consists in optimizing some criterion, i.e., a performance measure that allows to tune the model parameters to the data instance at hand. Such a model is then evaluated (used) by considering its generalization capability on unseen test patterns. Pattern recognition techniques can be grouped in two mainstream approaches: discriminative, such as support vector machines (Schölkopf and Smola 2002) and adaptive fuzzy inference systems (Sadeghian and Lavers 2011), and generative, like the hidden Markov models (Bicego et al. 2004) and the recently-developed deep convolutional neural networks (Sainath et al. 2014). Of course, also hybridized formulations exist.

Computational intelligence methods are typically designed relying on the assumption that the input space, $\mathcal{X}$, is essentially a subset of $\mathbb{R}^d$. When departing from the $\mathbb{R}^d$ pattern representation, theoretical and practical problems

arise, which are mostly due to the absence of an intuitive geometric interpretation of the data. However, many important applications can be tackled by representing patterns as "non-geometric" entities. For instance, it is possible to cite applications in document analysis (Bunke and Riesen 2011), solubility of *E. coli* proteome (Livi et al. 2015), bio-molecules recognition (Ceroni et al. 2007; Rupp and Schneider 2010), chemical structures generation (White and Wilson 2010), image analysis (Serratosa et al. 2013; Morales-González et al. 2014), and scene understanding (Brun et al. 2014). The availability of interesting datasets containing non-geometric data motivated the development of pattern recognition and soft computing techniques on such domains (Livi et al. 2014, 2015; Rossi et al. 2015; Fischer et al. 2015; Lange et al. 2015; Schleif 2014; Bianchi et al. 2015). Non-geometric patterns include data which are characterized by pairwise dissimilarities that are not metric; therefore they cannot be straightforwardly represented in a Euclidean space (Pękalska and Duin 2005). A particularly interesting instance of such non-geometric data is constituted by structured patterns. A labeled graph is the most general structured pattern that is conceivable, since it allows to characterize a pattern by describing the topological structure of its constituting elements (the vertices) through their relations (the edges) (Livi and Rizzi 2013). Both vertices and edges can be equipped with suitable labels, i.e., the specific attributes characterizing the elements and their relations. Sequences of generic objects, trees, and automata, for instance, can be though as particular instances of labeled graphs.

Figure 1 offers a schematic, visual representation of the typical stages involved in the use of CI methods for recognition purposes. By means of a training set (Tr in the figures), a model is synthesized and then used during the so-called testing stage by processing unseen data (denoted as Ts). Such a schematic organization is valid for both $\mathbb{R}^d$ (Fig. 1a) and non-geometric (Fig. 1b) data as well. However, in order to properly use standard CI methods in the case of non-geometric data, the input space must be processed with particular care. Notably, three mainstream approaches can be pursued (Livi et al. 2015): (i) using a suitable dissimilarity measure operating in the input space, (ii) using positive definite (PD) kernel functions, and (iii) embedding the input space in $\mathbb{R}^n$. The choice depends on the particular data-driven system adopted for the task at hand and on other factors, such as the computational complexity and the specific application setting. The first case is the most straightforward one, but its use is legitimate only if the data-driven system does not require a specific geometrical structure of the input space. In fact, a dissimilarity measure might not be metric; therefore also not Euclidean. The second case is a typical choice in the
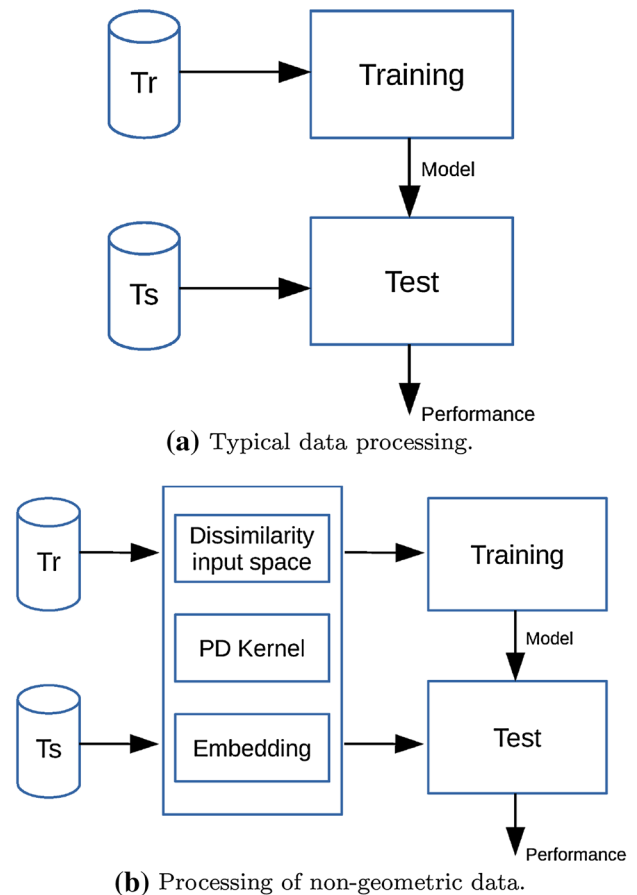

(a) Typical data processing.

(b) Processing of non-geometric data.

Fig. 1 Schematic representations of a data-driven inference system operating in $\mathbb{R}^d$ (**a**) and considering non-geometric (**b**) data, respectively

case of kernel methods—such as support vector machines. Positive definite kernels can be obtained only if the underlying geometry is Euclidean. However, if such a requirement does not hold, corrections techniques could be used to rectify the data (Livi et al. 2015). The last approach consists in mapping (i.e., embedding) the input data into a vector space, typically $\mathbb{R}^n$. In this way, conventional CI methods can be used without alterations.

3 Granular computing as a general data analysis framework

Granular computing (GrC) can be pictured as a general data analysis framework (or a data analysis paradigm) founded on IGs. Information granulation is a the basis of GrC: from the formation of sound IGs to their use in intelligent systems. IGs are the main mathematical constructs involved in the process of GrC. Several formalisms are available to implement IGs, such as fuzzy and rough

sets (Yao et al. 2013). Such mathematical frameworks offer solid bases on which designing IG models and their operations. However, an issue arises when we project our perspectives in a data-driven context, that is, when we try to extract (or synthesize) IGs from empirically observed data. A criterion to synthesize IGs from data is thus of utmost importance, since (i) it provides a way to formalize ideas under a common umbrella facilitating their practical and formal evaluations, and (ii) it lies at the basis of any consistent, formal theory. We suggest that a criterion for synthesizing IGs should play the same role as the one played by error functions in learning from data. A sound criterion for synthesizing IGs would lay the groundwork for conceiving a formal theory of GrC. Nonetheless, to date a general, unified, and consolidated theory of GrC, bridging model-based and data-driven perspectives in information granulation, is currently missing.

The integration of IG constructs and CI systems, such as pattern recognition and control systems, is nowadays well-established. For instance, granular neural networks offer an interesting example (Ding et al. 2014; Zhang et al. 2008; Ganivada et al. 2011; Song and Pedrycz 2013). Granular neural networks are basically extensions of typical artificial neural network architectures, which incorporate a mechanism of information granulation at the weights level or within the neuron model. In the first case, numerical weights modeling the synaptic connections of the network are realized in term of IGs—e.g., interval, fuzzy and rough sets. A particularly interesting consequence of this design choice is the fact that a granular neural network typically produces a granular output, hence consistent with the framework chosen for the IGs. Another interesting application of IGs can be cited in higher-order fuzzy inference systems (Biglarbegian et al. 2010; Gaxiola et al. 2014; Soto et al. 2014; Mendel 2014). Fuzzy inference systems and their extensions played a pivotal role in many applications in the last few decades. Higher-order fuzzy inference systems employ fuzzy sets of higher type instead of the original (type-1) fuzzy sets for handling the uncertainty of the input–output mapping and for performing the inference (e.g., rule composition). Clustering is another important research endeavor in which the GrC paradigm plays an important role (Tang and Zhu 2013; Linda and Manic 2012; Izakian et al. 2015). In fact, clustering algorithms are one of the most prominent example of techniques to generate IGs—technically, via the generation of a partition of the input data. Here IGs find a one-to-one mapping with clusters, which are typically endowed with some mathematical construct in order to offer a synthetic description of the data together with its characteristic uncertainty. IG constructs (mostly fuzzy sets) have been used also in problems of optimization and decision-making (Liang and Liao 2007; Kahraman et al. 2006; Pedrycz

2014; Wang et al. 2014a, b; Pedrycz and Bargiela 2012). In fact, both problems are typically affected by uncertainty at different levels: in the problem definition (e.g., constraints) or by considering the output (e.g., decision variables). It is worth citing the use of higher-order fuzzy constructs also in time series analysis (Chen and Tanuwijaya 2011; Huarng and Yu 2005), where either the time and amplitude (e.g., the time series realizations) domains are subjected to proper granulations. Finally, rough set theory found considerable application in many data analysis contexts. The rough set construct can be used to identify a reduced version of the original set of attributes pertaining to a decision system. As a consequence, rough sets found considerable application in feature selection and classification systems (Thangavel and Pethalakshmi 2009; Swiniarski and Skowron 2003; Foithong et al. 2012).

A founding prerequisite that an IG should satisfy is the capability of handling the uncertainty of the low-level entities that it aggregates. Klir (1995) states that "The nature of uncertainty depends on the mathematical theory within which problem situations are formalized". This fact suggests that the mathematical description of the data uncertainty pertaining a specific situation/process is not absolute, although it should be possible a reasonable and consistent mapping among the various theories. As a consequence, the specific setting on which IGs are defined affects in turn the way the data uncertainty is handled and therefore used in practice by an intelligent system operating through data granulation. Nonetheless, as postulated in Refs. (Livi and Sadeghian 2015; Livi and Rizzi 2015), the level of uncertainty conveyed by an IG defined according to some mathematical setting should be monotonically related to the uncertainty expressed by some other IG defined in a different setting. This suggests that, given some experimental evidence, the level of uncertainty is what should be preserved during data granulation, regardless of the formal setting used for defining IGs.

3.1 Information granules as data patterns

From a mathematical viewpoint, IGs are considered as formal constructs endowed with proper operations, such as intersection and union, to allow for symbolic manipulations. From a more operative perspective, instead, IGs typically play the role of computational components in some data-driven inference mechanism; as discussed in the previous sections. Nonetheless, more recently researchers (Ha et al. 2013; Guevara et al. 2014; Livi et al. 2013, 2014; Rizzi et al. 2013) realized that IGs could be considered also as a particular type of (non-geometric) patterns. This perspective opens the way to a multitude of future research works. For instance, it could be interesting to face typical pattern recognition problems, such as

clustering, classification, and function approximation, directly in the space of IGs. Technical issues involved in this process could be faced by exploiting the methods already developed for the non-geometric domains introduced in Sect. 2. Here, similarity measure for IGs (e.g., higher-order fuzzy sets, intuitionistic sets, and rough sets (Zhao et al. 2014; Chen and Chang 2015; Tahayori et al. 2015) could offer an important technical bridge between those fields. The process of data granulation implements an abstraction of the original data. This suggests that facing a data-driven problem in the space of IGs would required a different interpretation, offering thus also qualitatively different insights for the problem at hand.

3.2 General criteria for a justifiable data granulation

The quest for a general, sound, and justifiable criterion by which synthesizing IG from empirical evidence plays a pivotal role in GrC. IGs, in the data-driven setting, are obtained by means of an algorithmic procedure operating on some (typically, but not necessarily, non-granulated) input dataset. As previously stated, there are many mathematical models suitable for modeling IGs, like hyperboxes, (higher-order) fuzzy sets, shadowed sets, rough sets, and hybrid models (Pedrycz et al. 2008). All those models are framed in specific theories, having well-defined mathematical operations and descriptive measures. However, when facing (data-driven) problems involving the synthesis of IGs from a given dataset (experimental evidence), a sound and general criterion must be adopted. Such a criterion should be general, in the sense that it should work regardless of the specific IG model that is adopted. In fact, a general theory of GrC should not be conceived by focusing on specific mathematical formalisms for IGs. In addition, the criterion should be mathematically sound, meaning that it should admit a well-defined mathematical formulation, allowing thus for rigorous implementations, extensions, and validations. According to the perspectives offered in Sect. 2, we would be tempted to suggest that such a criterion should be applicable also regardless of the nature of the input data domain (e.g., numeric or not). In our opinion, such a criterion would provide the basic component for aiming toward a formal and unified theory of GrC, bridging both model-based and data-driven perspectives.

Despite the considerable effort recently devoted to the design of granulation procedures (algorithms for generating IGs) (Yao et al. 2013; Salehi et al. 2015) and formal GrC frameworks (Qian et al. 2014, 2015), to our knowledge it is possible to cite only two instances of such a criterion: the Principle of Justifiable Granularity (PJG) (Pedrycz and Homenda 2013) and the Principle of Uncertainty Level Preservation (PULP) (Livi and Sadeghian 2015).

3.2.1 The principle of justifiable granularity

The PJG (Pedrycz and Homenda 2013; Pedrycz 2011) has been developed as a guideline to form IGs from the available (experimental) input data. IGs generated following such a principle have to comply with two conflicting requirements: (i) justifiability and (ii) specificity. The first requirement insures that each IG would cover a sufficient portion of the experimental evidence. This means that a well-formed IG should not be too specialized. On the other hand, the second requirement provides a way to generate IGs that are not too dispersive, in the sense that an IG should come also with a well-defined semantics. Such two requirements taken together allow for a data-driven, user-centered, and problem-dependent synthesis of IGs from specific input datasets. The PJG itself is general—it has been successfully used to generate different IG types, including fuzzy sets and shadowed sets—and mathematically sound—it usually takes the form of an optimization problem. However, being designed in a user-oriented perspective, it is not conceived to directly offer a built-in mechanism to objectively evaluate the "quality" of the granulation itself. To this end, it is necessary to rely on external performance measures to quantify and judge over the quality of an IG.

3.2.2 The principle of uncertainty level preservation

Principle of uncertainty level preservation (Livi and Rizzi 2015; Livi and Sadeghian 2015) elaborates on a different perspective, by considering data granulation as a mapping between some input and output domain. PULP takes inspiration from the principles of uncertainty (Klir 1995), formulated by Klir few decades ago. A quantification of the uncertainty is considered in PULP as an invariant property to be preserved during the process of data granulation, i.e., when assigning an IG to a given input dataset. Uncertainty in PULP assumes the form of entropy expressions, exploiting the fact that the concept of entropy is well-developed in many mathematical settings, such as those of probability and fuzzy set theory. Therefore, synthesis of IGs is effectively casted in an information-theoretic framework, where the entropy measured for the input evidence is used as a guideline to form output IGs. Notably, the difference among the input–output entropy is considered as the granulation error, which needs to be minimized in order to reach a satisfactory result. PULP allows for a nonlinear relationship among the input–output entropy by considering a suitable monotonically increasing

function, bridging the two frameworks that define the input and output uncertainty quantifications, respectively.

Moving to mesoscopic descriptors, such as those provided by entropic characterizations, allows to conceived data granulation on a more abstract perspective. This fact has a number of benefits: (i) it automatically provides a way to quantify the quality of the granulation itself by objectively relying on the input–output uncertainty difference; (ii) it is applicable to any form of input data and IG formalism (at least to those where entropic functionals can be developed), and (iii) it allows to judge over the performance of different data granulation procedures operating in the same experimental conditions.

4 Concluding remarks

We would like to conclude this paper by posing a question: is granular computing an intrinsically experimental discipline? In other terms, is it possible to conceive an axiomatic theory of granular computing on which consistently develop both theoretical results and algorithmic solutions to perform data granulation and related operations? Of course, in the affirmative case, such a theory should be general, without focusing on specific models of information granules. A very important issue, as discussed in this paper, would be a criterion to bridge the model-based and data-driven perspectives of information granulation. That is to say, how should we perform data granulation? By following what criterion? We suggest that such a criterion would play an important role in the quest for a unified and sound theory of granular computing.

References

Ali MI, Davvaz B, Shabir M (2013) Some properties of generalized rough sets. Inf Sci 224:170–179. doi:10.1016/j.ins.2012.10.026

Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20(1):87–96. doi:10.1016/S0165-0114(86)80034-3

Bargiela A, Pedrycz W (2008) Toward a theory of granular computing for human-centered information processing. IEEE Trans Fuzzy Syst 16(2):320–330. doi:10.1109/TFUZZ.2007.905912

Bianchi FM, Livi L, Rizzi A (2015) Two density-based k-means initialization algorithms for non-metric data clustering. Pattern Anal Appl, pp 1–19. doi:10.1007/s10044-014-0440-4

Bicego M, Murino V, Figueiredo MAT (2004) Similarity-based classification of sequences using hidden Markov models. Pattern Recogn 37(12):2281–2291. doi:10.1016/j.patcog.2004.04.005

Biglarbegian M, Melek WW, Mendel JM (2010) On the stability of interval type-2 TSK fuzzy logic control systems. IEEE Trans Syst Man Cybern Part B Cybern 40(3):798–818. doi:10.1109/TSMCB.2009.2029986

Bonissone PP (1997) Soft computing: the convergence of emerging reasoning technologies. Soft Comput 1(1):6–18. doi:10.1007/s005000050002

Bronevich A, Klir GJ (2010) Measures of uncertainty for imprecise probabilities: an axiomatic approach. Int J Approx Reason 51(4):365–390. doi:10.1016/j.ijar.2009.11.003

Brun L, Saggese A, Vento M (2014) Dynamic scene understanding for behavior analysis based on string kernels. IEEE Trans Circuits Syst Video Technol 24(10):1669–1681. doi:10.1109/TCSVT.2014.2302521

Bunke H, Riesen K (2011) Recent advances in graph-based pattern recognition with applications in document analysis. Pattern Recogn 44(5):1057–1067. doi:10.1016/j.patcog.2010.11.015

Ceroni A, Costa F, Frasconi P (2007) Classification of small molecules by two-and three-dimensional decomposition kernels. Bioinformatics 23(16):2038–2045. doi:10.1093/bioinformatics/btm298

Chen Y, Wu K, Chen X, Tang C, Zhu Q (2014) An entropy-based uncertainty measurement approach in neighborhood systems. Inf Sci 279:239–250. doi:10.1016/j.ins.2014.03.117

Chen S-M, Chang C-H (2015) A novel similarity measure between Atanassov's intuitionistic fuzzy sets based on transformation techniques with applications to pattern recognition. Inf Sci 291:96–114. doi:10.1016/j.ins.2014.07.033

Chen S-M, Tanuwijaya K (2011) Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques. Expert Syst Appl 38(12):15425–15437. doi:10.1016/j.eswa.2011.06.019

Dai J, Tian H (2013) Entropy measures and granularity measures for set-valued information systems. Inf Sci 240:72–82. doi:10.1016/j.ins.2013.03.045

Ding S, Jia H, Chen J, Jin F (2014) Granular neural networks. Artif Intell Rev 41(3):373–384. doi:10.1007/s10462-012-9313-7

Engelbrecht AP (2007) Computational intelligence: an introduction. Wiley, Hoboken

Fischer A, Suen CY, Frinken V, Riesen K, Bunke H (2015) Approximation of graph edit distance based on Hausdorff matching. Pattern Recogn 48(2):331–343. doi:10.1016/j.patcog.2014.07.015

Foithong S, Pinngern O, Attachoo B (2012) Feature subset selection wrapper based on mutual information and rough sets. Expert Syst Appl 39(1):574–584. doi:10.1016/j.eswa.2011.07.048

Ganivada A, Dutta S, Pal SK (2011) Fuzzy rough granular neural networks, fuzzy granules, and classification. Theor Comput Sci 412(42):5834–5853. doi:10.1016/j.tcs.2011.05.038

Gaxiola F, Melin P, Valdez F, Castillo O (2014) Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction. Inf Sci 260:1–14. doi:10.1016/j.ins.2013.11.006

Guevara J, Hirata R, Canu S (2014) Positive definite kernel functions on fuzzy sets. In: Proceedings of the IEEE international conference on fuzzy systems, Beijing, China, pp 439–446. doi:10.1109/FUZZ-IEEE.2014.6891628

Ha M, Yang Y, Wang C (2013) A new support vector machine based on type-2 fuzzy samples. Soft Comput 17(11):2065–2074. doi:10.1007/s00500-013-1122-7

Han J, Lin TY (2010) Granular computing: models and applications. Int J Intell Syst 25(2):111–117. doi:10.1002/int.20390

Haykin S (2007) Neural networks: a comprehensive foundation. Prentice Hall PTR, Upper Saddle River

Huang B, Zhuang YL, Li HX (2013) Information granulation and uncertainty measures in interval-valued intuitionistic fuzzy information systems. Eur J Oper Res 231(1):162–170. doi:10.1016/j.ejor.2013.05.006

Huarng K, Yu H-K (2005) A type 2 fuzzy time series model for stock index forecasting. Phys A Stat Mech Appl 353:445–462. doi:10.1016/j.physa.2004.11.070

Izakian H, Pedrycz W, Jamal I (2015) Fuzzy clustering of time series data using dynamic time warping distance. Eng Appl Artif Intell 39:235–244. doi:10.1016/j.engappai.2014.12.015

Kahraman C, Ertay T, Büyüközkan G (2006) A fuzzy optimization model for QFD planning process using analytic network approach. Eur J Oper Res 171(2):390–411. doi:10.1016/j.ejor.2004.09.016

Klir GJ (1995) Principles of uncertainty: what are they? Why do we need them? Fuzzy Sets Syst 74(1):15–31. doi:10.1016/0165-0114(95)00032-G

Klir GJ (2006) Uncertainty and information: foundations of generalized information theory. Wiley-Interscience, Hoboken

Lange M, Biehl M, Villmann T (2015) Non-Euclidean principal component analysis by hebbian learning. Neurocomputing 147:107–119. doi:10.1016/j.neucom.2013.11.049

Liang RH, Liao JH (2007) A fuzzy-optimization approach for generation scheduling with wind and solar energy systems. IEEE Trans Power Syst 22(4):1665–1674. doi:10.1109/TPWRS.2007.907527

Linda O, Manic M (2012) General type-2 fuzzy c-means algorithm for uncertain fuzzy clustering. IEEE Trans Fuzzy Syst 20(5):883–897. doi:10.1109/TFUZZ.2012.2187453

Liu P, Li H (2004) Fuzzy neural network theory and application. World Scientific, Singapore

Livi L, Rizzi A, Sadeghian A (2014) Optimized dissimilarity space embedding for labeled graphs. Inf Sci 266:47–64. doi:10.1016/j.ins.2014.01.005

Livi L, Tahayori H, Sadeghian A, Rizzi A (2014) Distinguishability of interval type-2 fuzzy sets data by analyzing upper and lower membership functions. Appl Soft Comput 17:79–89. doi:10.1016/j.asoc.2013.12.020

Livi L, Rizzi A, Sadeghian A (2015) Granular modeling and computing approaches for intelligent analysis of non-geometric data. Appl Soft Comput 27:567–574. doi:10.1016/j.asoc.2014.08.072

Livi L, Giuliani A, Sadeghian A (2015) Characterization of graphs for protein structure modeling and recognition of solubility. Curr Bioinform. arXiv:1407.8033

Livi L, Rizzi A (2015) Modeling the uncertainty of a set of graphs using higher-order fuzzy sets. In: Sadeghian A, Tahayori H (eds) Frontiers of higher order fuzzy sets, pp 131–146. Springer, New York. doi:10.1007/978-1-4614-3442-9_7

Livi L, Rizzi A (2013) The graph matching problem. Pattern Anal Appl 16(3):253–283. doi:10.1007/s10044-012-0284-8

Livi L, Sadeghian A (2015) Data granulation by the principles of uncertainty. Pattern Recogn Lett. doi:10.1016/j.patrec.2015.04.008

Livi L, Sadeghian A, Pedrycz W (2015) Entropic one-class classifiers. IEEE Trans Neural Netw Learn Syst. doi:10.1109/TNNLS.2015.2418332

Livi L, Tahayori H, Sadeghian A, Rizzi A (2013) Aggregating α-planes for type-2 fuzzy set matching. In: Proceedings of the Joint IFSA World Congress and NAFIPS Annual Meeting, Edmonton, pp 860–865. doi:10.1109/IFSA-NAFIPS.2013.6608513

Melin P, Castillo O (2013) A review on the applications of type-2 fuzzy logic in classification and pattern recognition. Expert Syst Appl 40(13):5413–5423. doi:10.1016/j.eswa.2013.03.020

Mendel JM (2014) General type-2 fuzzy logic systems made simple: a tutorial. IEEE Trans Fuzzy Syst 22(5):1162–1182. doi:10.1109/TFUZZ.2013.2286414

Montes I, Pal NR, Janis V, Montes S (2015) Divergence measures for intuitionistic fuzzy sets. IEEE Trans Fuzzy Syst 23(2):444–456. doi:10.1109/TFUZZ.2014.2315654

Morales-González A, Acosta-Mendoza N, Gago-Alonso A, García-Reyes EB, Medina-Pagola JE (2014) A new proposal for graph-based image classification using frequent approximate sub-graphs. Pattern Recogn 47(1):169–177. doi:10.1016/j.patcog.2013.07.004

Nauck D, Klawonn F, Kruse R (1997) Foundations of neuro-fuzzy systems. Wiley, New York

Oh S-K, Kim W-D, Pedrycz W, Seo K (2014) Fuzzy radial basis function neural networks with information granulation and its parallel genetic optimization. Fuzzy Sets Syst 237:96–117. doi:10.1016/j.fss.2013.08.011

Pagola M, Lopez-Molina C, Fernandez J, Barrenechea E, Bustince H (2013) Interval type-2 fuzzy sets constructed from several membership functions: application to the fuzzy thresholding algorithm. IEEE Trans Fuzzy Syst 21(2):230–244. doi:10.1109/TFUZZ.2012.2209885

Pawlak Z (1982) Rough Sets. Int J Comput Inf Sci 11:341–356

Pedrycz W (1998) Shadowed sets: representing and processing fuzzy sets. IEEE Trans Syst Man Cybern Part B Cybern 28(1):103–109. doi:10.1109/3477.658584

Pedrycz W, Skowron A, Kreinovich V (2008) Handbook of granular computing. Wiley, New York

Pedrycz W (2011) Information granules and their use in schemes of knowledge management. Sci Iran 18(3):602–610. doi:10.1016/j.scient.2011.04.013

Pedrycz W (2013) Granular computing: analysis and design of intelligent systems. Taylor & Francis Group, Abingdon

Pedrycz W (2014) Allocation of information granularity in optimization and decision-making models: towards building the foundations of Granular Computing. Eur J Oper Res 232(1):137–145. doi:10.1016/j.ejor.2012.03.038

Pedrycz W, Succi G, Sillitti A, Iljazi J (2015) Data description: a general framework of information granules. Knowl Based Syst 80:98–108. doi:10.1016/j.knosys.2014.12.030

Pedrycz W, Bargiela A (2012) An optimization of allocation of information granularity in the interpretation of data structures: toward granular fuzzy clustering. IEEE Trans Syst Man Cybern Part B Cybern 42(3):582–590. doi:10.1109/TSMCB.2011.2170067

Pedrycz W, Chen S-M (2014) Information granularity, big data, and computational intelligence, vol 8. Springer International Publishing, Cham

Pedrycz W, Homenda W (2013) Building the fundamentals of granular computing: a principle of justifiable granularity. Appl Soft Comput 13(10):4209–4218. doi:10.1016/j.asoc.2013.06.017

Pękalska E, Duin RPW (2005) The dissimilarity representation for pattern recognition: foundations and applications. World Scientific, Singapore

Qian Y, Liang J, Dang C (2010) Incomplete multigranulation rough set. IEEE Trans Syst Man Cybern Part A Syst Hum 40(2):420–431. doi:10.1109/TSMCA.2009.2035436

Qian Y, Zhang H, Li F, Hu Q, Liang J (2014) Set-based granular computing: a lattice model. Int J Approx Reason 55(3):834–852. doi:10.1016/j.ijar.2013.11.001

Qian Y, Li Y, Liang J, Lin G, Dang C (2015) Fuzzy granular structure distance. IEEE Trans Fuzzy Syst PP(99):1. doi:10.1109/TFUZZ.2015.2417893

Rizzi A, Livi L, Tahayori H, Sadeghian A (2013) Matching general type-2 fuzzy sets by comparing the vertical slices. In: Proceedings of the Joint IFSA World Congress and NAFIPS Annual Meeting, pp 866–871. Edmonton. doi:10.1109/IFSA-NAFIPS.2013.6608514

Rossi L, Torsello A, Hancock ER (2015) Unfolding kernel embeddings of graphs: enhancing class separation through manifold learning. Pattern Recogn 48(11):3357–3370. doi:10.1016/j.patcog.2015.03.018

Rupp M, Schneider G (2010) Graph kernels for molecular similarity. Mol Inf 29(4):266–273. doi:10.1002/minf.200900080

Sadeghian A, Lavers JD (2011) Dynamic reconstruction of nonlinear v-i characteristic in electric arc furnaces using adaptive neuro-

fuzzy rule-based networks. Appl Soft Comput 11(1):1448–1456. doi:10.1016/j.asoc.2010.04.016

Sainath TN, Kingsbury B, Saon G, Soltau H, Mohamed A-R, Dahl G, Ramabhadran B (2014) Deep convolutional neural networks for large-scale speech tasks. Neural Netw 64:39–48. doi:10.1016/j.neunet.2014.08.005

Salehi S, Selamat A, Fujita H (2015) Systematic mapping study on granular computing. Knowl Based Syst 80:78–97. doi:10.1016/j.knosys.2015.02.018

Schleif F-M (2014) Generic probabilistic prototype based classification of vectorial and proximity data. Neurocomputing 154:208–216. doi:10.1016/j.neucom.2014.12.002

Schölkopf B, Smola AJ (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge

Serratosa F, Cortés X, Solé-Ribalta A (2013) Component retrieval based on a database of graphs for hand-written electronic-scheme digitalisation. Expert Syst Appl 40(7):2493–2502

Song M, Pedrycz W (2013) Granular neural networks: concepts and development schemes. IEEE Trans Neural Netw Learn Syst 24(4):542–553. doi:10.1109/TNNLS.2013.2237787

Soto J, Melin P, Castillo O (2014) Time series prediction using ensembles of ANFIS models with genetic optimization of interval type-2 and type-1 fuzzy integrators. Int J Hybrid Intell Syst 11(3):211–226. doi:10.3233/HIS-140196

Swiniarski RW, Skowron A (2003) Rough set methods in feature selection and recognition. Pattern Recogn Lett 24(6):833–849. doi:10.1016/S0167-8655(02)00196-4

Tahayori H, Livi L, Sadeghian A, Rizzi A (2015) Interval type-2 fuzzy set reconstruction based on fuzzy information-theoretic kernels. IEEE Trans Fuzzy Syst 23(4):1014–1029. doi:10.1109/TFUZZ.2014.2336673

Tang X-Q, Zhu P (2013) Hierarchical clustering problems and analysis of fuzzy proximity relation on granular space. IEEE Trans Fuzzy Syst 21(5):814–824. doi:10.1109/TFUZZ.2012.2230176

Thangavel K, Pethalakshmi A (2009) Dimensionality reduction based on rough set theory: a review. Appl Soft Comput 9(1):1–12. doi:10.1016/j.asoc.2008.05.006

Theodoridis S, Koutroumbas K (2008) Pattern Recognition, 4th edn. Elsevier/Academic Press, Waltham

Wagner C, Miller S, Garibaldi JM, Anderson DT, Havens TC (2015) From interval-valued data to general type-2 fuzzy sets. IEEE Trans Fuzzy Syst 23(2):248–269. doi:10.1109/TFUZZ.2014.2310734

Wagner C, Hagras H (2010) Toward general type-2 fuzzy logic systems based on zslices. IEEE Trans Fuzzy Syst 18(4):637–660. doi:10.1109/TFUZZ.2010.2045386

Wang S, Pedrycz W (2014a) Robust granular optimization: a structured approach for optimization under integrated uncertainty. IEEE Trans Fuzzy Syst. doi:10.1109/TFUZZ.2014.2360941

Wang S, Watada J, Pedrycz W (2014b) Granular robust Mean-CVaR feedstock flow planning for waste-to-energy systems under integrated uncertainty. IEEE Trans Cybern 44(10):1846–1857. doi:10.1109/TCYB.2013.2296500

White D, Wilson RC (2010) Generative models for chemical structures. J Chem Inf Model 50(7):1257–1274. doi:10.1021/ci9004089

Wu G-D, Zhu Z-W (2014) An enhanced discriminability recurrent fuzzy neural network for temporal classification problems. Fuzzy Sets Syst 237:47–62. doi:10.1016/j.fss.2013.05.007

Yao JT, Vasilakos AV, Pedrycz W (2013) Granular computing: perspectives and challenges. IEEE Trans Cybern 43(6):1977–1989. doi:10.1109/TSMCC.2012.2236648

Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353. doi:10.1016/S0019-9958(65)90241-X

Zhai D, Mendel JM (2011) Uncertainty measures for general type-2 fuzzy sets. Inf Sci 181(3):503–518. doi:10.1016/j.ins.2010.09.020

Zhang YQ, Jin B, Tang Y (2008) Granular neural networks with evolutionary interval learning. IEEE Trans Fuzzy Syst 16(2):309–319

Zhao T, Xiao J, Li Y, Deng X (2014) A new approach to similarity and inclusion measures between general type-2 fuzzy sets. Soft Comput 18(4):809–823. doi:10.1007/s00500-013-1101-z

Zhou S-M, Garibaldi JM, John RI, Chiclana F (2009) On constructing parsimonious type-2 fuzzy logic systems via influential rule selection. IEEE Trans Fuzzy Syst 17(3):654–667. doi:10.1109/TFUZZ.2008.928597

Zhu P, Wen Q (2012) Information-theoretic measures associated with rough set approximations. Inf Sci 212:33–43. doi:10.1016/j.ins.2012.05.014