Information Granules and Granular Models: Selected Design Investigations

Witold Pedrycz
Systems Research Institute
Polish Academy of Sciences
Warsaw, 01-447, Poland
and
Department of Electrical
& Computer Engineering
University of Alberta,
Edmonton T6R2G7AB, Canada
wpedrycz@ualberta.ca

Wladyslaw Homenda
the Faculty of Mathematics
and Information Science
Warsaw University of Technology
Warsaw, 00-662, Poland
and
the Faculty of Economics
and Informatics in Vilnius
University of Bialystok
LT-08221 Vilnius, Lithuania
homenda@mini.pw.edu.pl

Agnieszka Jastrzebska
the Faculty of Mathematics
and Information Science
Warsaw University of Technology
Warsaw, 00-662, Poland
A.Jastrzebska@mini.pw.edu.pl

Fusheng Yu
School of Mathematical Sciences
Beijing Normal University
Beijing 100875, China
yufusheng@bnu.edu.cn

***Abstract*— In the plethora of conceptual and algorithmic developments supporting system modeling, we encounter growing challenges associated with the complexity of systems, diversity of available data and a variety of requests imposed on the quality of the models. The accuracy of models is important. At the same time, the interpretability and explainability of models are equally important and of high practical relevance. We advocate that the level of abstraction at which models are constructed (and which could be flexibly adjusted), is conveniently realized through Granular Computing. Granular Computing is concerned with the development and processing information granules - formal entities that facilitate a way of organizing and representing knowledge about the available data and relationships existing there. This study identifies the principles of Granular Computing, shows how information granules are constructed and subsequently used in the realization of models.**

Keywords*— Granular Computing, information granules, fuzzy sets, design of information granules, clustering, principle of justifiable granularity, aggregation

I. INTRODUCTION

The apparent reliance on data and experimental evidence in system modeling, decision-making, pattern recognition, and control engineering, just to enumerate several representative spheres of interest, entails the centrality of data and emphasizes their paramount role in data science. To capture the essence of data, facilitate building their essential descriptors and reveal key relationships, as well as having all these faculties realized efficiently as well as deliver transparent, comprehensive, and user-oriented results, we advocate a genuine need for transforming data into information granules. In the realized setting, information granules become regarded as conceptually sound knowledge tidbits over which various models could be developed and utilized.

A tendency, which is being witnessed more visibly nowadays, concerns human centricity. Data science and big data revolve around a two-way efficient interaction with users. Users interact with data analytics processes meaning that the terms such as data quality, actionability, transparency are of relevance and are provided in advance. With this regard, information granules emerge as a sound conceptual and algorithmic vehicle owing to their way of delivering a more general view at data, ignoring irrelevant details and supporting a suitable level of abstraction aligned with the nature of the problem at hand.

The study is structured into 7 sections. We start with a brief discussion on information granules and a notion of information granularity (Section 2). The development of information granules is presented in Section 3; here we focus on the role of clustering and fuzzy clustering regarded as a conceptual and algorithmic prerequisite for the construction of information granules. The focus here is on the principle of justifiable granularity (Section 3.2). Further extensions of the principle are discussed in Section 4. The symbolic view at information granules associated with the linguistic summarization are covered in Section 5. In the sequel, in Section 6 discussed are granular models and the methodologies of their design.

II. INFORMATION GRANULES AND INFORMATION GRANULARITY

The framework of Granular Computing along with a diversity of its formal settings offers a critically needed conceptual and algorithmic environment. A suitable perspective built with the aid of information granules is advantageous in realizing a suitable level of abstraction in system modeling. It also becomes instrumental when forming sound and pragmatic problem-oriented tradeoffs among precision of results, their easiness of interpretation, value, and stability (where all of these aspects contribute vividly to the general notion of actionability). Information granules are intuitively appealing constructs, which play a pivotal role in human cognitive and decision-making activities [1-4]. We perceive complex phenomena by organizing existing knowledge along with available experimental evidence and structuring them in a form of some meaningful, semantically sound entities, which are central to all ensuing processes of describing the world, reasoning about the environment, and support decision-making activities. One can stress the diversity of formal frameworks of information granules, one can refer to sets, fuzzy sets, rough sets, shadowed sets, probabilistic sets, etc.

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

Information granules naturally emerge when dealing with data, including those coming in the form of data streams. The ultimate objective is to describe the underlying phenomenon in an easily understood way and at a certain level of abstraction. This requires that we use a vocabulary of commonly encountered terms (concepts) and discover relationships between them and reveal possible linkages among the underlying concepts.

One can ascertain that (a) information granules are the key components of knowledge representation and processing, (b) the level of granularity of information granules (their size, to be more descriptive) becomes crucial to the problem description and an overall strategy of problem-solving, (c) hierarchy of information granules supports an important aspect of perception of phenomena and deliver a tangible way of dealing with complexity by focusing on the most essential facets of the problem, (d) there is no universal level of granularity of information; commonly the size of granules is problem-oriented and user-dependent.

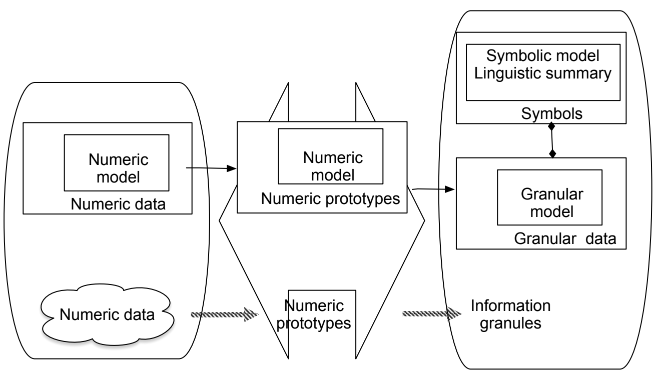

Fig. 1. Relationships among modeling environments and information granules: emphasized is the way of moving from experimental data to their representatives, information granules and linguistic summaries

In system modeling and models, information granules play at least two fundamental roles: (a) as building blocks using which a variety of models is built. The concept of granular models deals with models that establish mappings among information granules and realize tasks of prediction, classification and associations realized at a suitable level of abstraction implied by information granules. Granular models are constructed more efficiently than their numeric counterparts (the number of information granules is far smaller than the masses of numeric data), become more transparent and interpretable, (b) as a means to express the quality of the numeric models. In this case, granular models incorporate the mechanisms of granular processing and the parameters of the model are made granular following the optimal allocation of information granularity. With this regard, it is instructive to link the developments of information granules with how they support ways of system modeling as illustrated in Figure 1. Numeric data and numeric prototypes associate with numeric models. Granular prototypes give rise to granular models. The symbolic manifestation of information granules entails symbolic (qualitative) models; the symbols used there are well-grounded in virtue of the construction scheme supporting the buildup of information granules.

There are two clearly visible layers of processing. The one is concerned with the abstraction of available data: we proceed with numeric data (commonly acquired experimentally as a manifestation of the system under study), determine their numeric representatives (prototypes) and build information granules. In parallel, these activities give rise to particular processing realized in system modeling as portrayed at the upper portion of the figure.

III. Information Granules and Their Two-phase Development Process

Building information granules constitutes a central item on the agenda of Granular Computing with far-reaching implications on its applications. We present a way of moving from data to numeric representatives, information granules and then their linguistic summarization. The organization of the overall scheme and relationships among the resulting constructs are displayed in Figure 2.

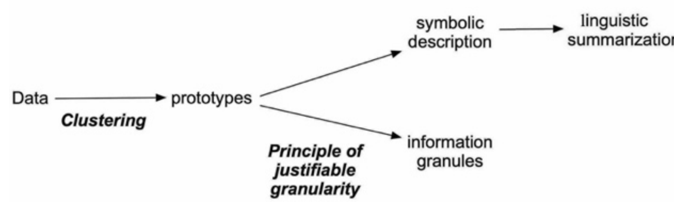

Fig. 2. From data to information granules and linguistic summarization

A. *Clustering as a prerequisite of information granules:*

Along with a truly remarkable diversity of detailed algorithms and optimization mechanisms of clustering, the paradigm itself delivers a viable prerequisite to the formation of information granules (associated with the ideas and terminology of fuzzy clustering, rough clustering, and others) and applies both to numeric data and information granules. Information granules built through clustering are predominantly data-driven, viz. clusters (either in the form of fuzzy sets, sets, or rough sets) are a manifestation of a structure encountered (discovered) in the data. Here we rely on a spectrum of objective function-based clustering, say *K*-Means or Fuzzy *C*-Means.

Numeric prototypes are formed through invoking clustering algorithms, which yield a partition matrix and a collection of the prototypes. Clustering realizes a certain process of abstraction producing a small number of the prototypes based on a large number of numeric data. Interestingly, clustering can be also completed in the feature space. In this situation, the algorithm returns a small collection of abstracted features (groups of features) that might be referred to as meta-features.

Two ways of generalization of numeric prototypes treated as key descriptors of data and manageable chunks of knowledge are considered: (i) symbolic and (ii) granular. In the symbolic generalization, one moves away from the numeric values of the prototypes and regards them as sequences of integer indexes (labels). Along this line, developed are concepts of (symbolic) stability and (symbolic) resemblance of data structures. The second generalization motivates the construction of information granules (*granular* prototypes), which arise as a direct quest for delivering a more comprehensive representation of the data than the one delivered through numeric entities. This entails that information granules (including their associated level of abstraction), have to be prudently formed to achieve the required quality of the granular model.

As a consequence, the performance evaluation embraces the following sound alternatives: (i) evaluation of representation capabilities of numeric prototypes, (ii) evaluation of representation capabilities of granular prototypes, and (iii) evaluation of the quality of the granular model.

1) Evaluation of representation capabilities of numeric prototypes: In the first situation, the representation capabilities of numeric prototypes are assessed with the aid of a so-called granulation-degranulation scheme yielding a certain reconstruction error. The essence of the scheme can be schematically portrayed as follows:

$\boldsymbol{x} \rightarrow$ internal representation $\rightarrow$ reconstruction

The formation of the internal representation is referred to as granulation (encoding) whereas the process of degranulation (decoding) can be sought as an inverse mechanism to the encoding scheme. In terms of detailed formulas, one encounters the following flow of computing:

a) encoding leading to the degrees of activation of information granules by input **x**, say $A_1(\mathbf{x})$, $A_2(\mathbf{x})$,.., $A_c(\mathbf{x})$ with

$$A_i(\boldsymbol{x}) = \frac{1}{\sum_{j=1}^{c} \left(\frac{\| \boldsymbol{x} - \boldsymbol{v}_i \|}{\| \boldsymbol{x} - \boldsymbol{v}_j \|} \right)^{2/(m-1)}}$$

in case the prototypes are developed with the use of the Fuzzy C-Means (FCM) clustering algorithm, the parameter *m* (>1) stands for the fuzzification coefficient and ||.|| denotes the Euclidean distance.

b) degranulation (decoding) producing a reconstruction of **x** via the following expression

$$\hat{\boldsymbol{x}} = \frac{\sum_{i=1}^{c} A_i^m(\boldsymbol{x}) \boldsymbol{v}_i}{\sum_{i=1}^{c} A_i^m(\boldsymbol{x})}$$

It is worth stressing that the above-stated formulas are a consequence of the underlying optimization problems. For any collection of numeric data, the reconstruction error is a sum of squared errors (distances) of the original data and their reconstructed versions.

B. *The principle of justifiable granularity*

The principle of justifiable granularity [5-7] guides a construction of an information granule based on available experimental evidence. In a nutshell, a resulting information granule becomes a summarization of data (viz. the available experimental evidence). The underlying rationale behind the principle is to deliver a concise and abstract characterization of the data such that (i) the produced granule is *justified* in light of the available experimental data, and (ii) the granule comes with a well-defined *semantics* meaning that it can be easily interpreted and becomes distinguishable from the others.

Formally speaking, these two intuitively appealing criteria are expressed by the criterion of coverage and the criterion of specificity. Coverage states how much data are positioned behind the constructed information granule. Put it differently – coverage quantifies an extent to which information granule is supported by available experimental evidence. Specificity, on the other hand, is concerned with the semantics of information granule stressing the meaning of the granule.

1) One-dimensional case: The definition of coverage and specificity requires formalization and this depends upon the formal nature of information granule to be formed. As an illustration, consider an interval form of information granule A. In case of intervals built on a basis of one-dimensional numeric data (evidence) $x_1, x_2, \ldots, x_N$, the coverage measure is associated with a count of the number of data embraced by *A*, namely

$$\mathrm{cov}(A) = \frac{1}{N} card\{x_k \mid x_k \in A\}$$

card (.) denotes the cardinality of *A*, viz. the number (count) of elements x_k belonging to *A*. In essence, coverage has a visible probabilistic flavor. The specificity of *A*, sp(*A*) is regarded as a decreasing function *g* of the size (length) of information granule. If the granule is composed of a single element, sp(*A*) attains the highest value and returns 1. If *A* is included in some other information granule *B*, then $sp(A) > sp(B)$. In a limit case if *A* is an entire space *sp*(*A*) returns zero. For an interval-valued information granule A =[*a*, *b*], a simple implementation of specificity with *g* being a linearly decreasing function comes as

$$sp(A) = g(length(A)) = 1 - \frac{|b-a|}{range}$$

where *range* stands for an entire space over which intervals are defined.

If we consider a fuzzy set as a formal setting for information granules, the definitions of coverage and specificity are reformulated to take into account the nature of membership functions admitting a notion of partial membership Here we invoke the fundamental representation theorem stating that any fuzzy set can be represented as a family of its α-cuts, namely

$$A(x) = \sup_{\alpha \in [0,1]} [\min(\alpha, A_\alpha(x))]$$

where

$$A_\alpha(x) = \{x \mid \mu(x) \geq \alpha\}$$

The supremum (sup) operation is taken over all values of α. In virtue of the theorem, we have any fuzzy set represented as a collection of sets. Having this in mind and considering (3) as a point of departure for constructs of sets (intervals), we have the following relationships

a) coverage

$$cov(A) = \int_X A(x)dx/N$$

where $\boldsymbol{X}$ is a space over which A is defined; moreover, one assumes that A can be integrated. The discrete version of the coverage expression comes in the form of the sum of membership degrees. If each data point is associated with some weight, the calculations of the coverage involve these values

$$cov(A)= \int_X w(x)A(x)dx / \int_X w(x)dx$$

b) specificity

$$sp(A) = \int_0^1 sp(A_\alpha)d\alpha$$

The criteria of coverage and specificity are in an obvious relationship, Figure 3. We are interested in forecasting temperature: the more specific the statement about the prediction is, the lower the likelihood of its satisfaction. To produce a meaningful prediction, a sound tradeoff between specificity and likelihood (coverage) needs to be established.

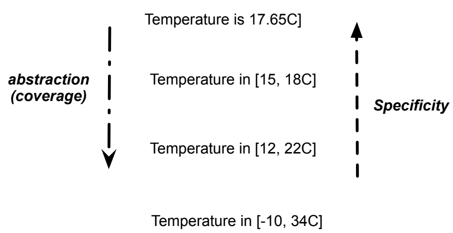

Fig. 3. Relationships between abstraction (coverage) and specificity of information granules of temperature

Let us introduce the following product of the criteria

$$V = cov(A)sp(A)$$

It is apparent that the coverage and specificity are in conflict; the increase in coverage associates with the drop in the specificity. Thus the desired solution is the one where the value of V attains its maximum.

The design of information granule is accomplished by maximizing the above product of coverage and specificity. Formally speaking consider that an information granule is described by a vector of parameters $\boldsymbol{p}$, $V(\boldsymbol{p})$. The principle of justifiable granularity gives to an information granule that maximizes V, $\boldsymbol{p}_{\text{opt}} = \arg_{\boldsymbol{p}} V(\boldsymbol{p})$.

To maximize the index V through adjusting the parameters of the information granule, two different strategies are encountered:

(i) a two-phase development is considered. First, a numeric representative (mean, median, mode, etc.) is determined. It can be regarded as an initial representation of the data. Next, the parameters of the information granule are optimized by maximizing V. For instance, in case of an interval, one has two bounds (a and b) to be determined. These two parameters are determined separately, viz. the a and b are formed by maximizing V(a) and V(b). The data used in the maximization of $V(b)$ involves the data larger than the numeric representative. Likewise, $V(a)$ is optimized on the basis of the data lower than this representative.

(ii) a single -phase procedure. Here all parameters of information granule are determined at the same time.

2) Multi-dimensional case: The results of clustering coming in the form of numeric prototypes $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_c$ can be further augmented by forming information granules giving rise to so-called granular prototypes. This can be regarded as a result of immediate usage of the principle of justifiable granularity and its algorithmic underpinning as elaborated earlier. Around the numeric prototype $\mathbf{v}_i$, one spans an information granule V_i, $V_i=(\mathbf{v}_i, \rho_i)$ whose optimal size is obtained as the result of the maximization of the well-known criterion

$$\rho_{i,opt} = \arg Max_{\rho_i}[\text{cov}(V_i)sp(V_i)]$$

where

$$\text{cov}(V_i) = \frac{1}{N} card\{\boldsymbol{x}_k \,|\, \|\boldsymbol{x}_k - \boldsymbol{v}_i\| \le \rho_i\} \quad \text{cov}(V_i) = 1 - \rho_i$$

assuming that we are concerned with normalized data. In the case of the FCM method, the data come with their membership grades (entries of the partition matrix). The coverage criterion is modified to reflect this. Let us introduce the following notation

$$\Omega_i = \{ \boldsymbol{x}_k \,|\, \|\boldsymbol{x}_k - \boldsymbol{v}_i\| \le \rho_i \}$$

Then the coverage is expressed in the form:

$$cov(V_i) = \frac{1}{N}\sum_{x_k \in \Omega_i} u_{ik}$$

which is a well-known σ-count used in fuzzy sets with u_{ik} being a membership value.

3) Representation aspects of granular prototypes: It is worth noting that having a collection of granular prototypes, one can conveniently assess their abilities to represent the original data (experimental evidence). The reconstruction problem, as outlined before for numeric data, can be formulated as follows: given $\mathbf{x}_k$, complete its granulation and de-granulation using the granular prototypes V_i, i=1, 2,.., c. The detailed computing generalizes the reconstruction process completed for the numeric prototypes and for given $\mathbf{x}$ yields a granular result $\hat{X} = (\hat{\boldsymbol{v}}, \hat{\rho})$ where

$$\hat{\boldsymbol{v}} = \frac{\sum_{i=1}^{c} A_i^m(x)\boldsymbol{v}_i}{\sum_{i=1}^{c} A_i^m(x)} \qquad \hat{\rho} = \frac{\sum_{i=1}^{c} A_i^m(x)\rho_i}{\sum_{i=1}^{c} A_i^m(x)}$$

The quality of reconstruction uses the coverage criterion formed with the aid of the Boolean predicate:

$$T(\boldsymbol{x}_k) = \begin{cases} 1 \text{ if } \|\hat{\boldsymbol{v}}_i - \boldsymbol{x}_k\| \le \hat{\rho} \\ 0, \text{ otherwise} \end{cases}$$

It is worth noting that in addition to the global measure of the quality of granular prototypes, one can associate with them their individual quality (taken as a product of the coverage and specificity computed in the formation of the corresponding information granule).

The principle of justifiable granularity highlights an important facet of elevation of the type of information granularity: the result of capturing a number of pieces of numeric experimental evidence comes as a single abstract entity - information granule. As various numeric data can be thought as information granule of type-0, the result becomes a single information granule of type-1. This is a general phenomenon of elevation of the type of information granularity. For instance, type-2 fuzzy sets [8] are constructs that result through the representation of a family of type-1 fuzzy sets. The increased level of abstraction is a direct consequence of the diversity present in the originally available granules. This elevation effect is of a general nature and can be emphasized by stating that when dealing with experimental evidence composed of information granules of type-*n*, the result becomes a single information granule of type (*n*+1).

As a way of constructing information granules, the principle of justifiable granularity exhibits a significant level of generality in two essential ways. First, given the underlying requirements of coverage and specificity, different formalisms of information granules can be engaged. Second, experimental evidence could be expressed as information granules articulated in different formalisms and on this basis, certain information granule is being formed

It is worth stressing that there is a striking difference between clustering and the principle of justifiable granularity. First, clustering leads to the formation of at least two information granules (clusters) whereas the principle of justifiable granularity produces a single information granule. Second, when positioning clustering and the principle vis-à-vis each other, the principle of justifiable granularity can be sought as a follow-up step facilitating an augmentation of the numeric representative of the cluster (such as e.g., a prototype) and yielding granular prototypes where the facet of information granularity is retained.

IV. Augmentation of the Design Process of Information Granules

So far, the principle of justifiable granularity presented is concerned with a generic scenario meaning that experimental evidence gives rise to a single information granule. Several conceptual augmentations are considered where several sources of auxiliary information are supplied:

A. Involvement of auxiliary variable

Typically, these could be some dependent variable one encounters in regression and classification problems. An information granule is built on a basis of experimental evidence gathered for some input variable and now the associated dependent variable is engaged. In the formulation of the principle of justifiable granularity, this additional information impacts a way in which the coverage is determined. In more detail, we discount the coverage; in its calculations, one has to take into account the nature of experimental evidence assessed on a basis of some external source of knowledge. In regression problems (continuous output/dependent variable), in the calculations of specificity, we consider the variability of the dependent variable y falling within the realm of A. More precisely, the value of coverage is discounted by taking this variability into consideration. In more detail, the modified value of coverage is expressed as

$$\text{cov}'(A) = \text{cov}(A)\exp(-\beta\sigma_y^2)$$

where σ is a standard deviation of the output values associated with the inputs being involved in the calculations of the original coverage cov(A). β is a certain calibration factor controlling an impact of the variability encountered in the output space. Obviously, the discount effect is noticeable, cov'(A)<cov(A).

In case of a classification problem in which p classes are involved, say $\omega = \{\omega_1\ \omega_2\ldots, \omega_p\}$, the coverage is again modified (discounted) by the diversity of the data embraced by the information granule where this diversity is quantified in the form of the entropy function $h(\omega)$

$$\text{cov}'(A) = \text{cov}(A)(1-h(\omega))$$

This expression penalizes the diversity of the data contributing to the information granule and not being homogeneous in terms of class membership. The higher the entropy, the lower the coverage cov'(A) reflecting the accumulated diversity of the data falling within the umbrella of A. If all data for which A has been formed belong to the same class, the entropy returns zero and the coverage is not reduced, cov'(A) = cov(A).

B. Adversarial information granules

In untargeted adversarial attacks [9,10], one considers $\boldsymbol{x}'$ such that it is close to the data coming from the training data $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$ and producing significantly different results than those reported for the neighboring data. The nature of the adversarial data $\boldsymbol{x}'$ can be quantified and generalized to the idea of the granular adversarial data. In light of the essence of the adversarial property, we determine $\boldsymbol{x}'$ such that it is close to $\boldsymbol{x}_k$ and $f(\boldsymbol{x}')$ is different from $f(\boldsymbol{x}_k)$ where $f(.)$ is a certain classifier or a model realizing this mapping f. $\boldsymbol{x}'$ is sought as an adversarial example.

The granular adversarial data centered around $\boldsymbol{x}'$ and denoted by $A(\boldsymbol{x}'; \rho)$ whose size ρ is the one which maximize the following ratio

$$V(\rho) = \frac{\sum_{\boldsymbol{x}_k:||\boldsymbol{x}_k-\boldsymbol{x}'||\leq n\rho^2} |f(\boldsymbol{x}_k) - f(\boldsymbol{x}')|}{\sum_{\boldsymbol{x}_k:||\boldsymbol{x}_k-\boldsymbol{x}'||\leq n\rho^2} ||\boldsymbol{x}_k - \boldsymbol{x}'||}$$

viz.

$$\rho_{\max} = \arg \text{Max}_\rho\ V(\rho)$$

where ||.|| is a certain distance function, say the Euclidean one.

Having the training data, one can assess the adversarial nature of individual datum by picking up $\boldsymbol{x}'=\boldsymbol{x}_k$ and determining the maximum of V and reporting the associated size (radius) ρ thus producing $A(\boldsymbol{x}_k; \rho_k)$. In this way, the data can be ranked with respect to their adversarial property by ordering the corresponding values of $V(\rho_k)$, and forming the resulting sequence starting from the highest value of this index.

V. Symbolic View at Information Granules and Their Symbolic Characterization and Summarization

Information granules are described through numeric parameters (or eventually granular parameters in case of information granules of higher type). There is an alternative view at a collection of information granules where we tend to move away from numeric details and instead look at the granules as symbols and engage them in further symbolic processing. Interestingly, symbolic processing is vividly manifested in Artificial Intelligence (AI). Consider that a collection of the prototypes has been generated as a result of clustering. The prototypes are projected on the individual variables (features) and their projections are ordered linearly. At the same time, the distinguishability of the prototypes is evaluated: if two projected prototypes are close to each other they are deemed indistinguishable and collapsed. The merging condition involves the distance between the two close prototypes: if $|v_i - v_{i+1}| < range/c\varepsilon$ then the prototypes are deemed indistinguishable. Here the *range* is the range of values assumed by the prototypes and ε is a certain threshold value less than 1. Once this phase has been completed, A_is are represented in a concise manner as sequences of indexes $I_i = (i_1, i_2, \ldots, i_{ni})$.

Linguistic summarization of numeric prototypes

Consider the symbolic representation of the prototypes A_1, A_2, …, A_c, namely $\boldsymbol{i}_1, \boldsymbol{i}_2, \ldots, \boldsymbol{i}_c$ where each $\boldsymbol{i}_k$ is a string of integer indexes where each index assumes values from 1 to *c*. The linguistic summarization of the prototype (or information granule) gives rise to the expressions such as *most* (attributes of the granule are *high*), at *most* 50% (attributes are *low*), etc. [11,12]. Generally speaking, the summarization is of the format

$$\tau(\text{attributes of granule is } \mu) = \lambda$$

j=1, 2,.., r. The optimized result of the summarization results from the maximization of the following expression

$$(i_0, j_0) = \arg\max_{i,j} \lambda_{ij}$$

VI. Granular Models

So far, the principle of justifiable granularity presented is concerned with a generic scenario meaning that experimental evidence gives rise to a single information granule. Several conceptual augmentations are considered where several sources of auxiliary information are supplied:

A. The concept

The paradigm shift implied by the engagement of information granules becomes manifested in several tangible ways including (i) a stronger dependence on data when building structure-free, user-oriented, and versatile models spanned over selected representatives of experimental data, (ii) emergence of models at various varying levels of abstraction (generality) being delivered by the specificity/generality of information granules, and (iii) building a collection of individual local models and supporting their efficient aggregation.

Here several main conceptually and algorithmically far-reaching avenues are emphasized. Notably, some of them have been studied to some extent in the past and several open up new directions worth investigating and pursuing. In what follows, we elaborate on them in more detail pointing at the relationships among them.

1) data → numeric models: This is a traditionally explored path being present in system modeling for decades. The original numeric data are used to build the model. There are several models, both linear and nonlinear exploiting various design technologies, estimation techniques and learning mechanisms associated with evaluation criteria where accuracy and interpretability are commonly exploited with the Occam razor principle assuming a central role. The precision of the model is an advantage however the realization of the model is impacted by the dimensionality of the data (making a realization of some models not feasible); questions of memorization and a lack of generalization abilities are also central to the design practices.

2) data → numeric prototypes: This path associates with the concise representation of data utilizing a small number of representatives (prototypes). The tasks falling within this scope are preliminary to data analytics problems. Various clustering algorithms constitute generic development vehicles using which the prototypes are built as a direct product of the grouping method.

3) data → numeric prototypes → symbolic prototypes: This alternative branches off to symbolic prototypes where on purpose we ignore the numeric details of the prototypes with intent to deliver a qualitative view at the information granules. Along this line, concepts such as symbolic (qualitative, i.e. not quantitaive) stability and qualitative resemblance of structure in data are established.

4) data → numeric prototypes → granular prototypes: This path augments the previous one by bringing the next phase in which the numeric prototypes are enriched by their granular counterparts. The granular prototypes are built in such a way so that they deliver a comprehensive description of the data. The principle of justifiable granularity helps quantify the quality of the granules as well as deliver a global view at the granular characterization of the data.

5) data → numeric prototypes → symbolic prototypes →qualitative modeling: The alternative envisioned here builds upon the one where symbolic prototypes are formed and subsequently used in the formation of qualitative models, viz. the models capturing qualitative dependencies among input and output variables. This coincides with the well-known subarea of AI known as qualitative modeling, see [13] with many applications [14-16].

6) data → numeric prototypes → granular prototypes →granular models: This path constitutes a direct extension of the previous one when granular prototypes are sought as a collection of high-level abstract data based on which a model is being constructed. In virtue of the granular data, we refer to such models as granular models.

B. Construction of granular models

There are two fundamental ways of constructing granular models:

1) stepwise development. One starts with a numeric model developed with the use of the existing methodology and algorithms and then elevate the numeric parameters of the models to their granular counterparts following the way outlined above. This design process dwells upon the existing models and in this way, one takes full advantage of the existing modeling practices. By the same token, one can envision that the granular model delivers a substantial augmentation of the existing models. In this sense, we can talk about granular neural networks, granular fuzzy rule-based models. In essence, the design is concerned with the transformation $a \xrightarrow{\epsilon} A = G(a)$ applied to the individual parameters where G is a certain formalism of information granulation. It is worth noticing that in the overall process, there are two performance indexes optimized: in the numeric model one usually considers the root mean squared error (RMSE) while in the granular augmentation of the model one invokes another performance index that takes into consideration the product of the coverage and specificity.

2) a single step design. One proceeds with the development of the granular model from scratch by designing granular parameters of the model. This process uses only a single performance index that is of interest to evaluate the quality of the granular result. Proceeding with the first design process presented above, the example presented in Figure 4 stresses a way in which granular models are formed.

The way of transforming (elevating) a numeric entity a to an information granule ***A*** is expressed formally as follows ***A*** = *G*(***a***). Here *G* stands for a formal setting of information granules. (say, intervals, fuzzy sets, etc.) and ε denotes a level of information granularity.

One can envision two among a number of ways of elevating a certain numeric parameter *a* into its granular counterpart.

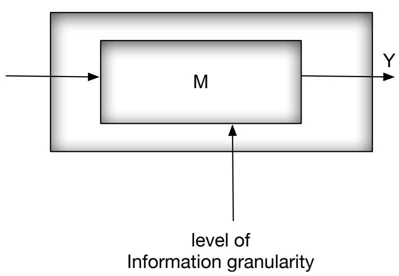

Fig. 4. From numeric to granular models by the development of granular parameters

One starts with a numeric (type-0) model $M(\boldsymbol{x}; \boldsymbol{a})$ developed on a basis of input-output data $\boldsymbol{D} = (\boldsymbol{x}_k, target_k)$. Then one elevates it to type-1 by optimizing a level of information granularity allocated to the numeric parameters ***a*** thus making them granular.

The way of transforming (elevating) a numeric entity a to an information granule ***A*** is expressed formally as follows ***A*** = *G*(***a***). Here *G* stands for a formal setting of information granules. (say, intervals, fuzzy sets, etc.) and ε denotes a level of information granularity. One can envision two among a number of ways of elevating a certain numeric parameter *a* into its granular counterpart. If *A* is an interval, its bounds are determined as

$$A = [\min(a(1-\varepsilon), a(1+\varepsilon)), \max((a(1-\varepsilon), a(1+\varepsilon))], \ \varepsilon \in [0,1]$$

Another option comes in the form

$$A = [\min(a/(1+\varepsilon), a(1+\varepsilon)), \max(a/(1+\varepsilon), a(1+\varepsilon))], \varepsilon \geq 0$$

If *A* is realized as a fuzzy set, one can regard the bounds of its support determined in the ways outlined above. Obviously, the higher the value of ε, the broader the result and higher likelihood of satisfying the coverage requirement. Higher values of ε yield lower specificity of the results. Following the principle of justifiable granularity, we maximize the product of coverage and specificity by choosing a value of ε.

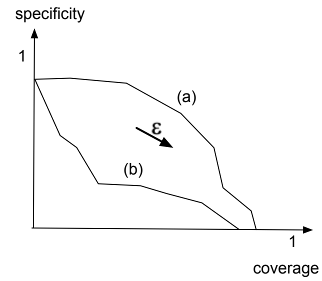

Fig. 5. Characteristics of granular models presented in the coverage-specificity coordinates

The performance of the granular model can be studied by analyzing the values of coverage and specificity for various values of the level of information granularity ε. Some plots of such relationships are presented in Figure 5. In general, as noted earlier, increasing the values of ε will result in higher coverage and lower values of specificity. An important is a quantification of the changes encountered here. By analyzing the pace of changes of the coverage versus the changes in the specificity, one can select a preferred value of ε as such beyond which the coverage does nor increase in a substantial way yet the specificity deteriorates significantly. One can refer to Figure 5 where both the curves (a) and (b) help identify the suitable values of ε. One can develop a global descriptor of the quality of the granular model by computing the area under the curve; the larger the area is, the better the overall quality of the model (quantified over all levels of information granularity) is. For instance, in Figure 5, the granular model (a) exhibits better performance than (b).

The level of information granularity allocated to all parameters of the model is the same. Different levels of information granularity can be assigned to individual parameters; an allocation of these levels could be optimized in such a way that the values of the performance index *V* becomes maximized whereas a balance of information granularity is retained. Formally, we consider the following optimization task

$$\text{Min}_{\boldsymbol{\varepsilon}} V \qquad \text{subject to constraints} \qquad \sum_{i=1}^{P} \varepsilon_i = P\varepsilon \quad \varepsilon_i \in [0,1]$$

where the vector of levels of information granularity is expressed as $\boldsymbol{\varepsilon} = [\varepsilon_1 \ \varepsilon_2 \ldots \ \varepsilon_P]$ and *P* stands for the number of parameters of the model. In virtue of the nature of this optimization problem, the use of evolutionary methods could be a viable option.

Granular models of higher type

The design of granular models of higher type is associated by admitting the results that are information granules of a higher type. For instance, in case of granular models of type-2, we envision the elevation of the type of information granularity

$a \xrightarrow{\epsilon} A$ (type-1) and $A \xrightarrow{\sigma} A\sim$(type-2) $\quad \sum_{i=1}^{P} \varepsilon_i = P\varepsilon \quad \varepsilon_i \geq 0$

In the design process, as before one starts with the development of the numeric model. The granular model of type-1 is obtained by optimizing the granular results with the use of the principle of justifiable granularity. The type-1 granular model cover some data ***D***; denote by ***D***' the data not covered by this granular model. Using *D*', we construct a granular model of type-2 again with the use of the principle of justifiable granularity, Fig. 5.

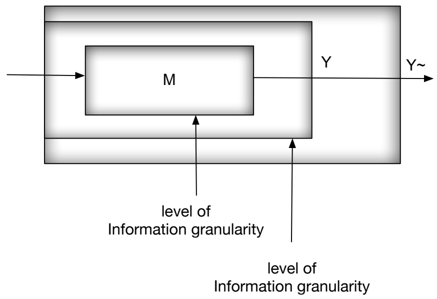

Fig. 5. Nesting granular models of higher type; from numeric to type-2 granular model

The formation of granular models of a higher type is inherently associated with the parameters of the model which are also information granules of higher type, refer to Fig. 6; the types of granular parameters correspond with the types of the granular models.

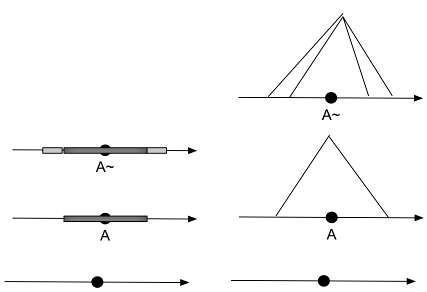

Fig. 6. From numeric to type-2 parameters of the models; shown are interval information granules and fuzzy sets

VII. GRANULAR MODELS

The study has offered a focused overview of the fundamentals of Granular Computing positioned in the context of advanced system modeling. We identified a multifaceted role of information granules as meaningful conceptual entities being formed at the required level of abstraction. It has been emphasized that information granules are not only reflective of the nature of the data (the principle of justifiable granularity highlights the reliance of granules on available experimental evidence) but can efficiently capture some auxiliary domain knowledge conveyed by the user and in this way reflect the human-centricity aspects of the investigations and enhances the actionability aspects of the results. The interpretation of information granules at the qualitative (linguistic) level and their emerging characteristics such as e.g., stability enhancement of the interpretability capabilities of the framework of processing information granules is another important aspect of data analytics that directly aligns with the requirements expressed by the user. Several key avenues of system modeling based on the principles of Granular Computing were highlighted; while some of them were subject of intensive studies, some other require further investigations.

By no means, the study is complete; instead, it can be regarded as a solid departure point identifying main directions of further far-reaching human-centric data analysis investigations. Several promising avenues are open that are well aligned with the current challenges of data analytics including the reconciliation of results realized in the presence of various sources of knowledge (models, results of analysis), hierarchies of findings, quantification of tradeoffs between accuracy and interpretability (transparency).

REFERENCES

[1] A.Bargiela, W. Pedrycz, ***Granular Computing: An Introduction***, Kluwer Academic Publishers, Dordrecht, 2003.

[2] A. Bargiela, W. Pedrycz, Toward a theory of Granular Computing for human-centered information processing, ***IEEE Transactions on Fuzzy Systems***, 16, 2, 2008, 320 – 330.

[3] L.A. Zadeh, Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic**,** ***Fuzzy Sets and Systems***, 90, 1997, 111-117.

[4] L.A. Zadeh, From computing with numbers to computing with words-from manipulation of measurements to manipulation of perceptions, ***IEEE Trans. on Circuits and Systems***, 45, 1999, 105-119.

[5] W. Pedrycz, W. Homenda, Building the fundamentals of granular computing: A principle of justifiable granularity, ***Applied Soft Computing***, 13, 2013, 4209-4218

[6] W. Pedrycz, ***Granular Computing***, CRC Press, Boca Raton, Fl, 2013.

[7] W. Pedrycz, Granular computing for data analytics: a manifesto of human-centric computing, ***IEEE /CAA Journal of Automatica Sinica***, 5, 2018, 1025-1034.

[8] J. M. Mendel, R. I. John, F. Liu, Interval Type-2 fuzzy logic systems made simple, ***IEEE Transactions on Fuzzy Systems***, 14, 2006,808-82

[9] A. Kurakin, I. Goodfellow, S. Bengio, Adversarial examples in the physical world. 2016**.** ***arXiv***:1607.02533.

[10] I. J. Goodfellow, J. Shlens, C. Szegedy C. Explaining and harnessing adversarial examples. 2014. ***arXiv***:1412.6572.

[11] J. Kacprzyk ; S. Zadrozny, Computing With Words is an implementable paradigm: fuzzy queries, linguistic data summaries, and natural-language generation, ***IEEE Transactions on Fuzzy Systems***, 18, 2010, 461-472.

[12] J. Kacprzyk ; R. R. Yager ; J. M. Merigo, Towards Human-centric aggregation via ordered weighted aggregation operators and linguistic data summaries: A new perspective on Zadeh's inspirations**,** ***IEEE Computational Intelligence Magazine***, 14, 2019, 15-30.

[13] K. Forbus. Qualitative process theory, ***Artificial Intelligence***, 24:85–168, 1984.

[14] W. Abou-Jaoudé, D. Thieffry, J. Feret, Formal derivation of qualitative dynamical models from biochemical networks, ***Biosystems***, 149, 2016, 70-112.

[15] N. Bolloju, Formulation of qualitative models using fuzzy logic, ***Decision Support Systems***, 17, 4, 1996, 275-298.

[16] W. Haider, J. Hu, J. Slay, B. P. Turnbull, Y. Xie, Generating realistic intrusion detection system dataset based on fuzzy qualitative modeling, ***Journal of Network and Computer Applications***, 87, 2017, 185-192.