
GRANULAR COMPUTING AS AN EMERGING PARADIGM OF INFORMATION PROCESSING

1. 1 INTRODUCTORY COMMENTS

This book is about granular computing, its fundamentals, methodologies, algorithms and applications. In a nutshell, granular computing, as the name itself stipulates, deals with representing information in the form of some aggregates (embracing a number of individual entities) and their processing. Information granules are everywhere. They are central to processes of abstraction guiding our intellectual pursuits. Without any exaggeration one can state that processing at the level of information granules is a predominant feature of knowledge-intensive systems. This chapter serves as a concise and gentle introduction to the subject. First, we introduce the notion of information granularity through a discussion of several illustrative examples that come from commonly visible and representative areas of engineering and science. Second, we elaborate on a number of formal models of information granules and their processing. Along this line comes a discussion on the conceptual and algorithmic aspects of information granules such as their granularity, usefulness, communication and interoperability between various platforms of granular computing.

We stress that the primary intent of this chapter is to elaborate on the fundamentals and put the entire area of granular computing in a certain global perspective and prepare a stage for more detailed analysis and synthesis.

1. 2 INFORMATION GRANULES ARE EVERYWHERE

Information granules, as the name itself stipulates, are collections of entities, usually originating at the numeric level, that are arranged together due to their similarity, functional adjacency, indistinguishability, coherency or alike (Pedrycz, 2001;

Bargiela, 2001; Pedrycz and Bargiela, 2002, Zadeh, 1979, 1997; Zadeh and Kacprzyk, 1999; Pedrycz and Vukovich, 1999; Pedrycz and Smith, 1999, Pedrycz, Smith, Bargiela, 2001). Information granules as all abstraction of our reality are aimed at building efficient and user-centered views of the external world and supporting and facilitating our perception of the surrounding physical and virtual world.

Let us elaborate on some representative areas with which information granulation is inherently associated.

Spatial Granulation: Image Processing and GIS

From its very inception, image processing has been confronted with a challenging goal of building intelligent systems that are capable of understanding and describing images. This challenge is still with us and considering a rapid pace of developments at the WWW involving indexing visual information we may anticipate further quests.

Image processing naturally splits into two main and overlapping levels of processing. The lower end of the processing deals with image segmentation, edge detection, noise removal, etc. At the higher end of abstraction, we are interested in image description and interpretation. Here the level of detail (or the level of abstraction) depends on the task we have to handle. Images perceived by humans are full of information granules. An image of any landscape consists of trees, houses, roads, lakes, shrubs, etc. They are spatially distributed and this distribution is an important factor in describing the content of the image. Interestingly, all these objects are generic information granules. In many cases there are no clear boundaries between them (say forests and marshes).

Spatial granulation is central to all GIS (Geographical Information Systems) processing. Maps forms hierarchies of abstractions and granulation realizes processes of abstraction. When establishing a coarse view of the world, we deal with large information granules: continents, countries, and oceans. We are concerned with abstractions at a high level. When more details are required, we move down to regions, provinces, states, seas, etc. All minute details are revealed to us when moving down to specific maps of towns, lakes, forests, etc. The level of information granulation depends heavily on the task at hand and the need of the decision-making process.

Temporal Granulation

Time is an omnipresent variable in all human endeavors. As such its granulation is of paramount importance and happens everywhere (Dyreson et al., 2000). Granulation of time incorporate the cultural, legal, business orientation of the

designer. The granularity of time depends upon the application. On one hand, we deal with strategic planning when plans are developed based on a horizon of 10-15 years and the meaningful granules span over several years. Short term plans operate at the level of months and quarters. We talk about days when dealing with date of birth. We use very refined information granules when talking about clock cycles of a computer. In this sense, information granules carry a well-defined semantics. Collection of information granules are referred to as calendars. The hierarchy such the temporal information granules becomes evident. It becomes evident that the level of information granularity implies here an effect of indeterminacy (obviously, it occurs in any process of information granulation). For instance, when concerned about the days treated as the most detailed information granules and expressing there a birth date, it means that we encounter a “don’t know when” effect at the level of more detailed information granules (hours or minutes): we know that the person was born sometime during the given day; the precise hour or minute are not known.

Some other examples of information granules are briefly highlighted below

- in any computer system we granulate memory resources by subscribing to the notion of pages of memory as its basic operational chunks (then we may consider various swapping techniques to facilitate an efficient access to individual data items)
- In describing any problem, we tend to shy away from numbers. Instead, we tend using aggregates and building rules (*if-then statements*) that dwell on them.
- We live in an inherently analog world. Computers, by tradition and technology, perform processing in a digital world. Digitization of this nature (that dwells on set theory - interval analysis) is an example of information granulation
- All mechanisms of data compression are examples of information granulation that is carried in a certain sense

In all the examples shown above, we dealt with mechanisms of building information granules. The process of constructing information granules is referred to as information granulation. As already indicated, no matter how this granulation proceeds and what fundamental technology it involves, there are several essential factors that drive all pursuits of information granulation. Several interesting and important observations need to be made here

- A need to split the problem into a sequence of more manageable and smaller subtasks. Here granulation serves as an efficient vehicle to modularize the problem. The primary intent is to reduce an overall computing effort by exercising a fundamental principle of “*divide and conquer*” that permeates a majority of problems of system design and analysis
- A need to comprehend the problem and provide with a better insight into its essence rather than get buried in all unnecessary details. In this sense,

granulation serves as an abstraction mechanism that reduces an entire conceptual burden. As a matter of fact, by changing the “size” of the information granules, we can hide or reveal a certain amount of details one intends to deal with during a certain design phase.

- Information granulation and the ensuing processing is *human-centric* meaning that the user, designer, developer are in the center of all these endeavours. Information granules do *not* exist as tangible physical entities but they are conceptual entities that emerge in the ocean of information. Their emergence is implied by the needs of humans in a continuous quest to abstraction, summarization, condensation of information. Information granulation supports communication at different levels: between humans, humans and computers, computers and computers (the latter trend being fully reflected in terms of autonomous agents). It is needless to say that in the world of information and its various manifestations including cyberspace, processing of information granules (granular computing) becomes a necessity.

We can easily recognize that these factors occur quite ostensibly in the general categories of the problems discussed before. In all cases the abstraction and its realization in the setting of information granules becomes apparent.

On a more detailed and application – driven note, we can state that information granules may arise as a phenomenon of inherent nonuniqueness associated with the problem at hand. As a simple example, one can resort himself to any inverse problem; the type of characteristics involved (as the functions may be non-invertible) gives rise to relations and as a result, a collection of information granules rather than single numeric quantities. Dropping some input variable in a model may also lead to the same effect of granular information.

As we can observe, the concept of granular computing tends to permeate a number of significant endeavors. The reason is quite straightforward. Granular computing as opposed to numeric computing is *knowledge-oriented*. Numeric computing is *data-oriented*. Undoubtedly, knowledge-inclined processing arises as a cornerstone of data mining, rule-based models, intelligent databases, hierarchical and supervisory control, just to name a few representative examples.

The long-lasting tradition of computing using some specific information granules is a visible testimony that some specific versions of granular computing are omnipresent indeed. As a matter of fact, as we will discuss in depth, a digital – to-analog transformation leading to *digital* computing in *analog* world is just a highly representative (albeit quite specific) instance of granular computing. By tradition (and the associated technology dominant at that time), we have embarked on the digital world of computing. To interact with the continuous (analog) world, we use set-based granulation (more specifically, interval-valued granulation). This specific type of granulation comes under the name of analog-to-digital conversion.

1.3 FORMAL MODELS OF INFORMATION GRANULES

The diversity of the formal means used for information granulation and further processing of the resulting information granules has a common denominator. All of these environments share the same research agenda of the common goal to address the fundamentals of granular computing.

The process of granulation and the nature of information granules implies a certain formalism that seems to be the most suited to capture the problem at hand. Intuitively, we note a difference when building a granule of sport cars and a granule of a forest in some GIS system. In the first case the granule exhibits definite boundaries: an element (car) is either in or out. We deal with a set theory as a suitable conceptual and algorithmic framework. In the second case the situation is radically different as the term “forest” does not exhibit clearly defined boundaries: an element shown in the map may belong to the granule to some degree. Intuitively, we need a different formalism to capture the nature of such information granules.

There are a number of formal frameworks in which information granules are built. They are well-known, thoroughly investigated and coming with a vast array of applications:

- Set theory and interval analysis
- Fuzzy sets
- Rough sets
- Shadowed sets
- Probabilistic sets and probability-based granular constructs
- Higher-level granular constructs

What is worth noting, though, is a fact that most of them were developed independently without any significant interaction occurring between them.

From the general point of view, information granules defined in some space \mathbf{X} can be treated as a mapping

$$A: \mathbf{X} \rightarrow G(\mathbf{X})$$

where A is an information granule of interest. G denotes a formal framework of information granules. These could be sets in which case we use a notation $P(\mathbf{X})$, fuzzy sets with the notation $F(\mathbf{X})$, rough sets $R(\mathbf{X})$ (Polkowski and Skowron, 1998; Lin and Cercone, 1997), shadowed sets ($S(\mathbf{X})$) and alike. When referring to A , we always specify the framework of granulation in which A has been defined (so we say that A is an interval, fuzzy set or shadowed set).

Information granules come with underlying rules describing syntax and semantics. The semantics addresses the meaning conveyed by an information granule. As being a result of some abstraction, its meaning is well-defined and practically relevant. The syntax results directly from the formal environment in which the granules are formed. For instance, operations of aggregation such as union, intersection and others (negation, dilution, concentration, etc.) are defined as a part of the formalism being considered.

1.4 CONCEPTUAL ASPECTS OF INFORMATION GRANULES

In spite of the diversity of the formal frameworks, information granules can be described in a fairly general form meaning that there are some general characteristics that are common across all these platforms.

Size of Information Granules and their Relevance

The question as to the definition of the “size”, “capacity” or “dimension” of the information granule is of primordial interest. How to measure granularity of the constructed information granules? How to relate this granularity with computational complexity? From the intuitive point of view, the size of the granule describes its specificity. We say how specific the granule is and how many details it embraces. The more elements we identify as belonging to the granule, the more abstract and general it becomes. Its further application implies that any model in which such information granules are used can address the problem at the corresponding level of generality (specificity). The notion of cardinality (again expressed in the pertinent language of sets, fuzzy sets, etc.) is the one commonly used. Computing the cardinality is about enumerating (counting) the number of elements in the information granule. In more detail, we may quantify granularity through an integral of the form

$$\text{Card}(A) = \int_{\mathbf{x}} A(\mathbf{x}) d\mathbf{x}$$

where A is an information granule under consideration (being more precise, we describe A in the form pertinent to the assumed formal framework of granulation such as sets, fuzzy sets, rough sets, etc.). The higher the cardinality, the higher the abstraction of the granule and the lower its specificity. Obviously, the above expression is the simplest possible and one can think of functionals of the form

$$\int_{\mathbf{x}} F(A(\mathbf{x})) d\mathbf{x}$$

where F is a certain monotonically increasing transformation of A .

Obviously in all cases we assume that such integrals do make sense.

Usefulness of Information Granules

The level of information granularity is implied by the problem in which such granules are used. We have already pointed out that information granules can be treated as conceptual building blocks with the use of which we perceive and describe the problem as well as plan some interaction with the external world (such as planning through control or decision-making or pursuing various prediction tasks). The type of description and interaction dictates the level of granularity; the most relevant (useful) level becomes selected. There could be other reasons for choosing a certain level of granularity such as e.g., a computational effort that usually is directly tied up with the size of information granules. In this sense, we regard information granules as a usefully vehicle of carry out efficient computing. All in all, one can portray this matter of usefulness of information granules versus their level of granularity is illustrated in Figure 1. We stress that the usefulness in this figure is meant in some general way as discussed above. It is noticeable that such usefulness can vary quite significantly depending upon the problem at hand: in Figure 1 (a) we witness a case where with the increasing level of granularity the decline in the usefulness level is quite limited. Figure 1 (b) alludes to the situation in which the increase in the granularity level (where we start using more detailed granules) leads to quite a substantial deterioration of the usefulness (it could well be that this is a result of excessive computing effort or too detailed information used in guidance of high-end decision-making processes).

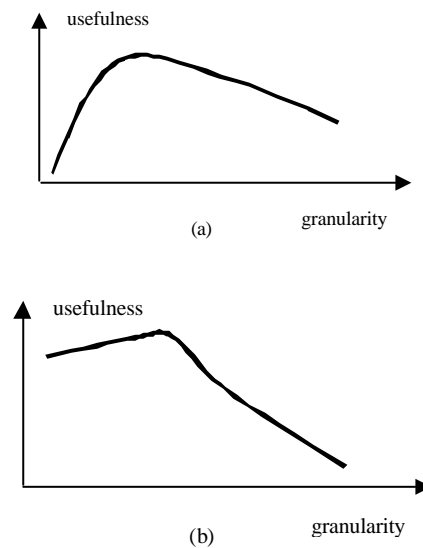


Figure 1. Usefulness of information granules as a function of their granularity: shown are two different profiles of usefulness.

1.5 DEFINING A GRANULAR WORLD

Once we have decided upon the use of some specific formal framework, we usually define a vocabulary of granular terms that are next viewed as a frame of reference. They are just a collection of information granules with well defined semantics that we may also regard as some conceptual landmarks or so-called reference information granules. For instance, when talking about traffic on a highway, we use granules such as low, medium, high, very high speed that are further used to describe traffic patterns. Similarly, we can define a collection of some temporal information granules defined in the time domain, say morning rush, lunch time, afternoon rush, evening hours, etc. These concepts help us form a general structure of a granular world – an environment of information granules that supports all processes of information granulation, information processing and information exchange (that supports various form of interaction with the external environment). We will be building this world step by step. The family of reference information granules (a frame of reference) $\mathbf{A} = \{A_1, A_2, \dots, A_c\}$ is an important component of the granular world. So far, we have discussed the formal framework of information granules G and the reference information granules (frame of reference). Now we put them together in the definition of the granular world

$$\mathbf{G} = \langle \mathbf{X}, G, \mathbf{A} \dots \rangle \quad (1)$$

(the dots indicate that there are yet some more components to be defined). The syntax of operations in \mathbf{G} is completely implied by the formal framework of information granulation G .

The frame of reference helps us express any information granules in the language of its elements, that is X being defined in \mathbf{X} can be described in terms of A_i s. For instance, when dealing with some speed, say $X = \text{about } 80 \text{ km/hr}$, we can express this particular granule in terms of \mathbf{A} that consists of the granules such as { low speed, medium speed, high speed, very high speed }.

The frames of reference could come at different levels of granularity. For instance, one may have $\mathbf{B} = \{B_1, B_2, \dots, B_p\}$ where “p” is substantially higher than “c” granules existing in the previous frame of reference. This new frame of reference implies a new granular world, $\mathbf{G}' = \langle \mathbf{X}, G, \mathbf{B} \dots \rangle$ Obviously, if we change the formal mechanism used to describe information granules, we end up with a new granular world. In the case of \mathbf{G} and \mathbf{G}' we talk about a (granular) hierarchy of the granular worlds; because of the way in which \mathbf{A} and \mathbf{B} have been formed, we say that \mathbf{G}' is a refinement of \mathbf{G} (or put it differently \mathbf{G} is an abstraction of \mathbf{G}').

Examples of relationships between granular worlds are shown in Figure 2. Note that while some of them are ordered in a linear way (because of a certain granularity of the frames of references), some others cannot be compared by having different formalism of information granulation.

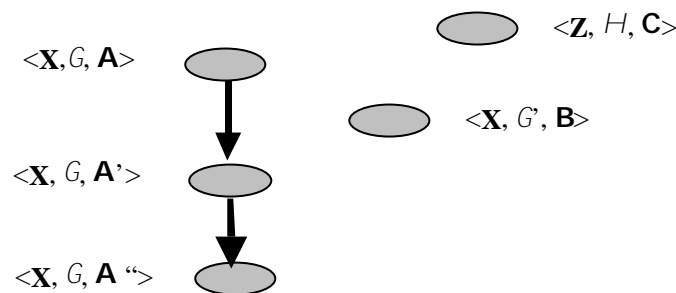


Figure 2. Relationships between granular worlds; note that some of them form a hierarchy.

1. 6 GRANULAR COMPUTING: AN INFORMATION PROCESSING PYRAMID

In granular computing we operate on information granules. As we have already noticed, information granules arise at different levels of granularity. Depending upon the problem at hand, we usually group granules of similar “size” (that is granularity) together in a single layer. If more detailed (and computationally intensive) processing is required, smaller information granules are sought. Then these granules are arranged in another layer. In total, the arrangement of this nature gives rise to the information pyramid. As portrayed schematically in Figure 3, in granular processing we encounter a number of conceptual and algorithmic layers indexed by the “size” of information granules.

Information granularity implies the usage of various techniques that are relevant for the specific level of granularity. Alluding to system modeling, we can refine Figure 4 by associating the layers of the information processing pyramid with the pertinent most commonly used classes of processing and resulting models

- at the lowest level we are concerned with numeric processing. This is a domain completely overwhelmed by numeric models such as differential equations, regression models, neural networks, etc.
- at the intermediate level we encounter larger information granules (viz. those embracing more individual elements)
- the highest level can be solely devoted to symbol-based processing and as such invokes well-known concepts of finite state machines, bond graphs, Petri nets,

qualitative simulation, etc. Note that some of these classes emerge at the intermediate level of information granularity and at that level their conceptual and symbolic fabric is usually augmented with some numeric component. It is worth stressing that the lowered granularity (higher abstraction) helps embark on models that involve logic and algebraic methods thus becoming more transparent.

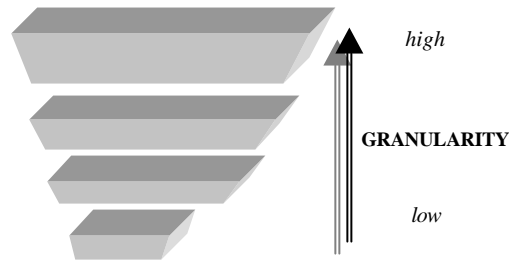


Figure 3. An information-processing pyramid (the respective layers are indexed by the corresponding level of information granularity that is granularity of the granular worlds involved).

The general characteristics of the principle of granular computing can be enumerated as shown in Table 1.

<i>Allow for multiple abstraction levels (granularity levels)</i>
<i>Allow for several methods of traversing various levels of hierarchy (encoding – decoding mechanisms)</i>
<i>Allow for nonhomogeneous methods (differential or difference equations, Petri nets, finite state machines)</i>

Table 1. The fundamental features of granular computing.

With the emergence of granular models, arise new fundamental questions as to evaluation of such constructs. Any evaluation criterion (viz. a performance index) needs to be associated with the granularity of information granules. In particular, if the model has been constructed in the setting $\langle \mathbf{X}, G, \mathbf{A}' \rangle$ then its evaluation in the environment $\langle \mathbf{X}, G, \mathbf{A}'' \rangle$ where \mathbf{A}'' is more specific than \mathbf{A}' may be excessively demanding. It is likely that the evaluation (and model testing) should be completed for the collection of information granules of the size that is the same and similar to those being used in the design of the model. In a similar vein, we can discuss the use of the granular model for new granular data (prediction problem) and an acceptable

level of granularity it can handle. This leads to several questions of a fundamental nature. Is the model developed with the use of “large” information granules useful when more specific results are required? It is apparent that when forming information granules, the contributing elements lose their identity that is essentially a non-recoverable process. Now, how this could effect the results of computing involving larger information granules (viz. those of lower granularity)? If we want to recover the details, how efficient could be our attempt? What are the limits of this reconstruction? These aspects boil down to the mechanisms of encoding and decoding granular information that will be discussed in the next section and becomes a part of a more general problem of communication between granular worlds.

1.7 COMMUNICATION BETWEEN GRANULAR WORLDS

Granular worlds rarely exist and operate independently without any interaction with the environment (that could be a physical world or some other granular world). Typically, we can consider various agents each of them endowed with some granular world. The agents interact between themselves and this manifests in some form of collaboration or competition. As each agent comes with its own environment of granular computing $\langle X, G, A \rangle$, $\langle Y, G', B \rangle$, $\langle Z, Q, C \rangle$. To allow for any communication, one has to assure that there are some mechanisms that help agents interact with. Schematically, the communication mechanisms can be shown as a certain layer developed around the agent, Figure 4.

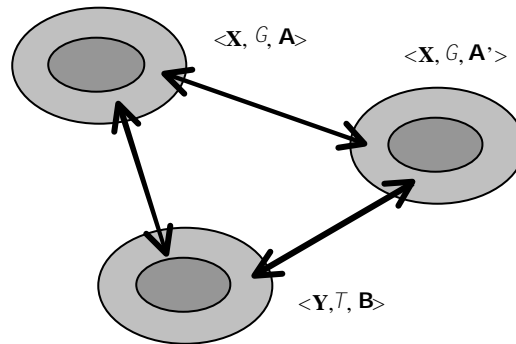


Figure 4. Collaboration between granular worlds; the mechanisms of interaction are displayed as an auxiliary processing layer around the agents.

As a consequence, the formal definition of the granular world needs to be augmented by the communication mechanisms; we add them as the family of communication procedures C ,

$$G = \langle X, G, A, C \rangle \quad (2)$$

where we mean that **C** may consists of a variety of constructs that help communicate (collaborate, compete, interact with the granular worlds based on different formal schemes of information granules).

In general, we distinguish between two main categories of the communication tasks. The first one involves two granular worlds built around the same formalism of information granules that is we are concerned with $\langle X, G, A \rangle$ and $\langle Y, G, B \rangle$; the granularity of **A** and **B** can differ quite substantially. The second category of the tasks in which we do not impose any constraint on the formalism of the granular information. To elaborate on the communication mechanisms in more detail and show what they really entail, we discuss two examples: the first one arising in the context of a traversal of the information pyramid of the models we showed before and the communication between two interval-based granular environments.

Fundamental Issues of Traversing Information Pyramid: Encoding and Decoding

Granular computing supports modeling activities carried out at various levels of information granularity, refer again to Figure 5.

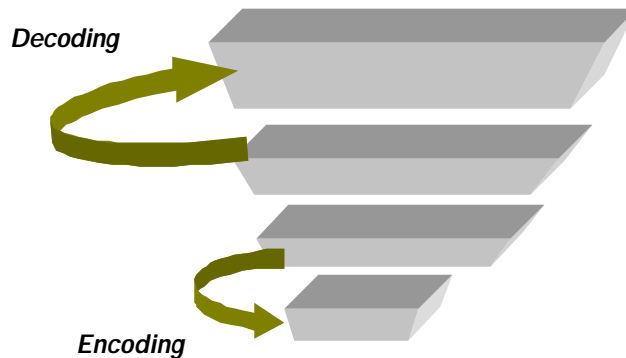


Figure 5. Decoding and encoding information granules as a vehicle of traversing the information pyramid.

The ability to traverse through the layers characterized by different sizes of information granules is one of the dominant features of the modeling pursuits discussed in this framework. Each modeling layer indexed by the assumed level of granularity, comes with its own repertoire of modeling techniques. For instance, for the highest level of information granularity, viz. numeric data, we are dealing with differential equations and regression models as basic vehicles of system modeling.

Commonly used neural networks fall under the same category. When moving towards nonnumeric layer where some information granules of lower granularity are formed, we encounter a diversity of models such as Petri nets, finite state machines, bond graphs, constraint-based, etc. (Boros et al., 2000; Harris and Brown, 1993; Kandel, 1986; Kasabov, 1996; Pedrycz, 1997; Zadeh and Kacprzyk, 1999). Depending on the specific form of granulation, we subsequently allude to fuzzy Petri nets, probabilistic Petri nets, etc.

The layers communicate between themselves. They receive data from other layers, complete computing (processing) and return the results to some other layers. These communication mechanisms are referred to as encoding and decoding, respectively. The role of the encoder is to transform the input information entering the given layer. The objective of the decoder is to convert the information granules produced by the given layer into the format acceptable by the destination layer. Depending on the problem at hand and the formalism of information granulation being used, a specific naming comes into play.

The general formulation of the encoding – decoding problem can be delineated as follows, see Figure 5: develop encoding (Enc) and associated decoding (Dec) algorithms such that the following relationship is satisfied

$$\text{Dec}(\text{Enc}(X)) = X$$

for all information granules $(X) \in G(X)$ that are defined in a certain formal framework of information granulation and for a broad range of sizes of the information granules involved. In a limit case numeric granules are also included. Note, however, that the decoding-encoding scheme could be very demanding and one may not be able to meet the equality. More practically, we request that the design of these transformation should minimize the associated transformation error meaning that we are interested in minimizing the expression involving the distance $\|.\|$ between the original information granule and its transformation

$$\|\text{Dec}(\text{Enc}(X)) - X\| \rightarrow \text{Min}$$

over a given range of granularity of X 's involved there and for a fixed granulation environment. As a matter of fact, the above minimization problem is not trivial. The resulting solution may very much depend upon the size of the information granules exploited in a granular world and the level of granularity of the information granule under discussion (X). Intuitively, if these levels of granularity are very much distinct, the distance between X and its decoded version, $\text{Dec}(\text{Enc}(X))$ could vary quite substantially.

The A/D and D/A conversions form an interesting (yet highly specific) illustration to the formulation of the problem given above, see Figure 6. It should be stressed that

in this case the granulation process assumes a well-known version of *discretization* (let us emphasize that granulation subsumes this scheme as a particular case. Essentially, we are confined to the set-based formalism). We get:

A/D: $\text{Enc}(X) : X = \{x\} \in \mathbf{R} \rightarrow X \in P(\mathbf{R})$ (the resulting granules are intervals in \mathbf{R} ; depending how the intervals are formed, one encounters either uniform quantization or a non-uniform one)

D/A: $\text{Dec}(X) : X \in P(\mathbf{R}) \rightarrow X = \{x'\} \in \mathbf{R}$ (usually a quantization error occurs so we never obtain the original numeric entity, $x \neq x'$).

The A/D and D/A conversions can be revisited and generalized in the framework of fuzzy sets, $F(\mathbf{X})$. This leads to the following formulation of the problem

A/D: $\text{Enc}(X) : X = \{x\} \in \mathbf{R} \rightarrow X \in F(\mathbf{R})$ (the resulting granules are fuzzy sets in \mathbf{R} ; depending how they are formed, one encounters either uniform quantization or a non-uniform linguistic discretization of \mathbf{X})

D/A: $\text{Dec}(X) : X \in P(\mathbf{R}) \rightarrow X = \{x'\} \in \mathbf{R}$ (usually a quantization error it can be avoided by selecting a proper family of fuzzy sets. The zero error occurs for the triangular fuzzy sets with $\frac{1}{2}$ overlap between successive membership functions).

In fuzzy controllers (Harris and Brown., 1993), the process of converting numeric data into the format accepted by the inference engine is called *fuzzification*. This is the name used for the encoding mechanism. The decoding is referred to as a *defuzzification* scheme.

One may also envision also a mixed form of information granules, namely they may originate from different formal environments of information granulation.

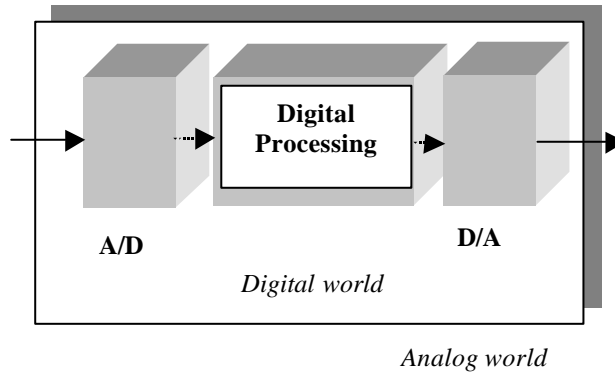


Figure 6. Digital processing as an example of commonly encountered granular computing; note a role of A/D and D/A converters utilized as the encoding and decoding modules.

Interoperability Between Different Formal Platforms of Information Granules

Various models of information granules and granulation processes themselves are crucial in the realization of interoperability when dealing with various platforms of granular computing. We illustrate this important concept in the setting of data mining or collaboration between autonomous agents. Information granules, no matter what formal framework they are supported by, are used as front and back end interfaces of the data mining computing machine. The need for the studies of the hybrid models of information granules arises when we are faced with an issue of interoperability between various tasks or subsystems of data mining that could be realized in various frameworks of granular computing. As an example, consider a situation visualized in Figure 7 (a). One data mining task, say T_1 is realized in the setting of information granules in the setting G_1 . The other one is developed in the granular environment G_2 . The results of the first task need to be communicated to the second module. This inherently gives rise to concept of the hybrid models of information granularity. For instance, assume that G_2 dwells on set theory. Now if G_1 generates the results in the form of fuzzy sets, this type of communication gives rise to fuzzy rough sets. Interestingly, even though G_1 and G_2 could exploit the same formalism of granular information, the communication between these two modules produces rough sets. This arises as a result of a certain level of granularity of data. As visualized in Figure 7 (b), "X is A" is a result of passing a message to the second task. This, in turn, invokes the representation of A in terms of the family of sets. As a consequence, even though we have sets at both ends, the representation of A emerges as a rough set. Put it in a different way: rough sets are just the outcome of the communication at the granular level. In more detail, refer to Figure 8, X is transformed into the following form

$$X \in P(\mathbf{X}) \Rightarrow X^* \in R(\mathbf{X})$$

with P and R being the families of sets and rough sets defined in \mathbf{X} . The lower and upper bound of X^* is expressed as

$$X_* = \{A_4\} \quad X^* = \{A_3, A_4, A_5\}$$

One can justify the origin and usage of some other hybrid models in the same way.

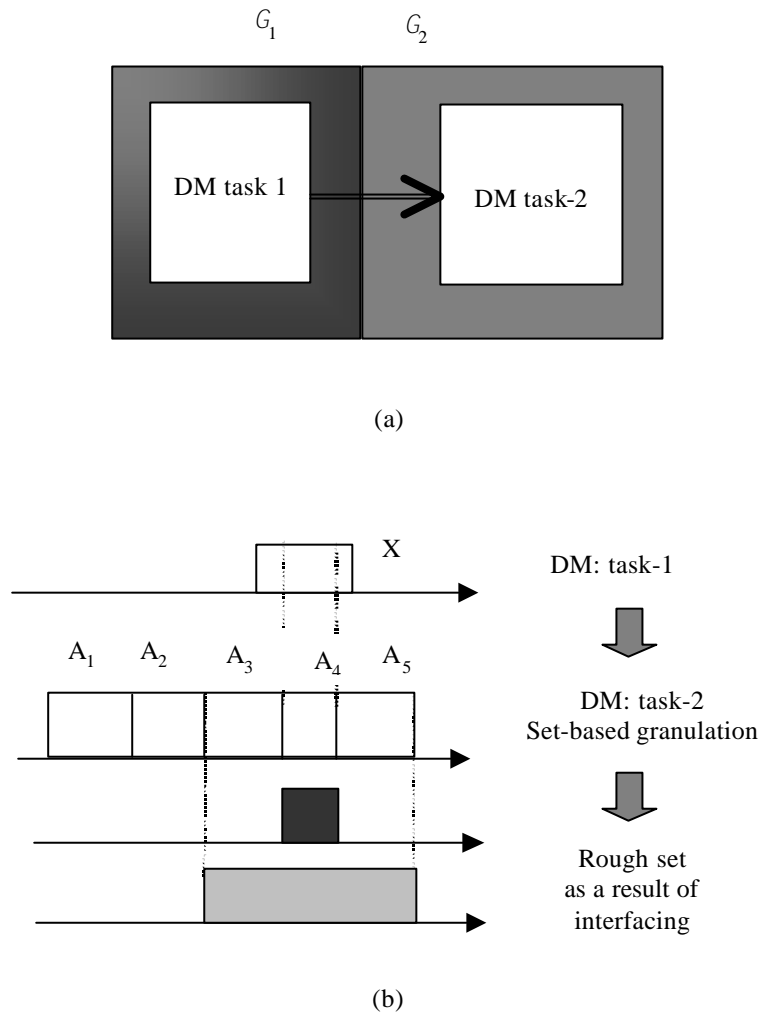


Figure 7. Two data mining tasks realized with the aid of different formalisms of information granulation: (a) a general scheme of communication, and (b) rough sets arising as an effect of communication between the two DM tasks accomplished in the granular setting implemented by sets.

1. 8 CONCLUSIONS

We have discussed the fundamentals of granular computing viewed as a new unified paradigm of processing information granules. Granular computing subsumes commonly encountered numeric processing as its special (limit) case.

The research agenda of granular computing includes a series of key and well-defined methodological and algorithmic issues

- Construction of information granules. This process deals both with the selection of the formal framework of information granulation and detailed estimation procedure producing information granules. The latter dwells on the usage of the setting in which the granules are constructed.
- Characterization of dimension (granularity) of information granules. This task is crucial as providing us with a better insight as to the essence of the granulation process and its implications both at the level of the methodology of the design of the ensuing granular model as well as its usage.
- The development of the encoding and decoding mechanisms. These are essential to the functioning of any granular architecture. The encoding and decoding schemes are essential to the performance of granular computing. Interestingly, the essence of information compatibility expressed in terms of its granularity is inherently related with granular computing and nonexistent within other environments.
- The issues of interoperability are crucial to the design of systems operating within the realm of various formalisms of information granularity.

This chapter should be treated as a bird-eye view of the rapidly growing research area. It concentrates on the methodology, attempts to identify the common features and help put the existing and somewhat scattered approaches under the same conceptual and algorithmic umbrella.

REFERENCES

- Bargiela, A. (2001) Interval and ellipsoidal uncertainty in water system state estimation, in: *Granular Computing*, (Pedrycz, W., ed.), Physica Verlag, 23-57.
- Bargiela, A., Pedrycz, W., Hirota, K. (2002), Logic-based granular prototyping, *Computers Software and Applications Conference, COMPSAC 2002*, Oxford, August 2002.
- Boros, E., Hammer, P.L., Ibaraki, T., Kogan, A., Mayoraz, E., Muchnik, I. (2000), An implementation of logical analysis of data, *IEEE Trans. on Knowledge and Data Engineering*, 12(2), 292-306.
- Dyreson, C.E., Evans, W.S., Lin, H., Snodgrass R. (2000), Efficiently supporting temporal granularities, *IEEE Trans. on Knowledge and Data Engineering*, 12(4), 568-587.
- Harris, C.J., Moore, C.G., Brown, M. (1993), *Intelligent Control - Aspects of Fuzzy Logic and Neural Nets*, World Scientific, Singapore.

- Kandel, A. (1986), *Fuzzy Mathematical Techniques with Applications*, Addison-Wesley.
- Kasabov, N. (1996), *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*, MIT Press, Cambridge, MA.
- Lin, T.Y., Cercone N. eds.(1997), *Rough Sets and Data Mining - Analysis of Imperfect Data*, Kluwer Academic Publishers, Boston.
- Pedrycz, W. (1997), *Computational Intelligence: An Introduction*, CRC Press, Boca Raton.
- Pedrycz, W., Smith, M. H. (1999), Granular correlation analysis in data mining, *Proc. 18th Int Conf of the North American Fuzzy Information Processing Society (NAFIPS)*, New York, June 1-12, 715-719.
- Pedrycz, W., Smith M.H., Bargiela, A. (2000), Granular clustering: A granular signature of data, *Proc. 19th Int. (IEEE) Conf. NAFIPS'2000*, Atlanta, July 2000, 69-73.
- Pedrycz, W., Vukovich, G. (1999), Quantification of fuzzy mappings: a relevance of rule-based architectures, *Proc. 18th Int Conf of the North American Fuzzy Information Processing Society (NAFIPS)*, New York, June 1-12, 105-109.
- Pedrycz W., ed. (2001), *Granular Computing: An Emerging Paradigm*, Physica-Verlag.
- Pedrycz, W., Bargiela, A. (2002), Granular clustering: A granular signature of data, *IEEE Trans. on Systems Man and Cybernetics*, Vol 32, No. 2, 212-224.
- Polkowski, L., Skowron, A. eds (1998), *Rough Sets in Knowledge Discovery*, vol.I: Methodology and Applications, Physica-Verlag, Heidelberg.
- Zadeh, L. A. (1979), Fuzzy sets and information granularity, In: M.M. Gupta, R.K. Ragade, R.R. Yager, eds., *Advances in Fuzzy Set Theory and Applications*, North Holland, Amsterdam, 3-18.
- Zadeh, L. A. (1997), Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, **90**, 111-117.
- Zadeh, L. A., Kacprzyk, J. (1999), *Computing with Words in Information/Intelligent Systems*, vol. 1-2, Physica-Verlag, Heidelberg.