Granular Computing for Human-Centred Systems Modelling

Grant: GR/R10707/01; **Investigators:** A. Bargiela & W. Pedrycz

Background/Context

The purpose of this research programme, as declared in the original proposal, was to investigate the foundations of the computer-based information processing in the context of human-centred engineering systems. The investigation was carried out against the background of a broad spectrum of supporting research in the Simulation and Modelling group at the Nottingham Trent University and the Computer Engineering group at the University of Alberta. As such, the research carried out within this project was specifically intended to focus on integration and systematisation of insights gained through other projects undertaken in the two groups and in other research centres. The main thesis tested in this research programme was that granular computing provides a human-centred, powerful conceptual and algorithmic framework that is capable of simplifying systems' descriptions and providing quantification of the confidence levels associated with system models.

The origins of the idea of Granular Computing can be traced back to the pioneering contribution of Prof. Zadeh entitled "Fuzzy Sets and Information Granularity", published in a book Advances in Fuzzy Set Theory and Applications in 1979. However, it was not until the subsequent contribution from Prof. Zadeh, entitled "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", published in Fuzzy Sets and Systems, 90, 111-127, 1997, that the idea of Granular Computing became accepted as a mainstream paradigm of Computational Intelligence. Subsequently, information granulation and information granules became a preferred vocabulary for describing numerous application domains. Well-known ideas of rule-based systems dwell inherently on information granules. Qualitative modelling, being one of the leading threads of AI, operates on a level of information granules. Multi-tier architectures and hierarchical systems (such as those encountered in control engineering), planning and scheduling systems all exploit information granularity. We also utilize information granules when it comes to functionality granulation, reusability of information and efficient ways of developing underlying information infrastructures.

For the first time, GC brings all of these formalisms and methodologies together, treats them uniformly and becomes the fundamental computing paradigm. GC establishes a sound research agenda that promotes synergies between the already well-established technologies of sets (intervals), fuzzy sets and rough sets. Is Granular Computing a totally new pursuit? The answer is: yes and no. GC has inherited a long tradition and specialized research agenda of the contributing technologies. On the other hand, Granular Computing opens up entirely new research avenues and promotes an innovative and holistic view of the fundamental mechanisms of information representation and processing.

Key Advances and Supporting Methodology

The main achievement of this research programme is gathering together and systematising the fundamentals, methodologies, algorithms and representative applications of Granular Computing. 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 abstractions of our reality are aimed at building efficient and user-centred views of the external world and supporting and facilitating our perception of the surrounding physical and virtual world. This is well appreciated by looking at some representative areas with which information granulation is inherently associated.

<u>*Spatial Granulation: Image Processing and GIS*</u>

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 spacially 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. Using the framework of granular computing we can identify the third, less obvious level of image processing that of semmantical analysis. This becomes aparrent when looking at the fine arts images that convey information that is not srtictly visual but that is prompted by visual clues. Artists have long realised that the works of art can be used to encode artistic messages and the messages are subsequently decoded by individual viewers using ontology of concepts and symbols. Somewhat surprisingly, we have dis covered a very natural bridge between computer information processing and communication of artistic messages. To emphasise this synergy we have coined a term of *Granular Art* and have subjected the granular interpretation of photographic images to the scrutiny of the artists (Exchibition in the Art Gallery "*Extravagance*" , Bargiela, 2002).

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

Granulation of time incorporates the cultural, legal, commercial, etc. 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 which clearly occurs in any process of information granulation.

Formal Models of Information Granules

The research undertaken here investigated several formal frameworks in which information granules can be built. These are:

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

It has been noted that these formalisms were developed independently and that they overlook significant interaction occurring between them. From the general point of view, information granules defined in some space **X** can be treated as a mapping

$$A: \mathbf{X} \longrightarrow \mathsf{G}(\mathbf{X})$$

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 $\mathsf{P}(\mathbf{X})$, fuzzy sets with the notation $\mathsf{F}(\mathbf{X})$, rough sets $\mathsf{R}(\mathbf{X})$ (Polkowski and Skowron, 1998; Lin and Cercone, 1997), shadowed sets ($\mathsf{S}(\mathbf{X})$) 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). The research has identified that the formal framework of sets, fuzzy sets and rough sets provides a sufficiently expressive environment for the description of a broad spectrum of real-life applications (Bargiela, Pedrycz, 2002).

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 (Bargiela, Pedrycz, 2002, Pedrycz, Bargiela, 2002)

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

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 $\int_X F(A(\mathbf{x}))d\mathbf{x}$ 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 usefuly 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).

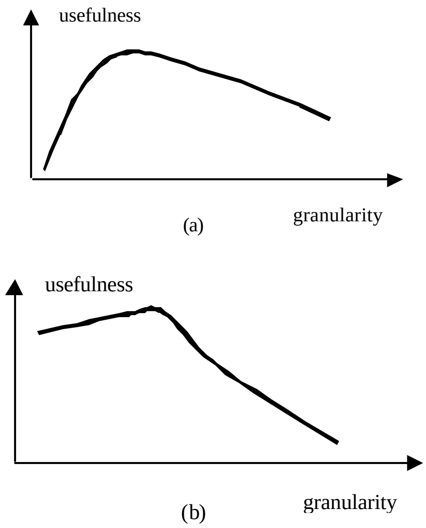

Figure 1. Usefulness of information granules as a function of their granularity: shown are two different profiles of of usefulness.

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. The family of reference information granules (a frame of reference) $\mathsf{A} = \{A_1, A_2, \ldots, A_c\}$ 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 (Pedrycz, 1997, Bargiela, Pedrycz, 2002)

$$G = < \mathbf{X}, G, A\ldots > \quad (1)$$

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

The frames of reference could come at different levels of granularity. For instance, one may have B = $\{B_1, B_2, \ldots, B_p\}$ 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, G' = < **X**, G, B…> Obvioulsly, if we change the formal mechanism used to describe information granules, we end up with a new granular world. In the case of G and G' we talk about a (granular) hierarchy of the granular worlds; because of the way in which A and B have been formed, we say that G ' is a refinement of G (or put it differently G is an abstraction of 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.

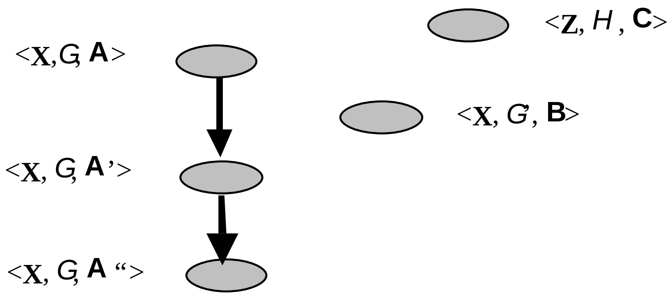

Figure 2. Relationships between granular worlds; note that some of them form a hierarchy.

<u>*Communication Between Granular Worlds*</u>

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 enowed with some granular world. The agents interact between themselves and this manifests in some form of collaboration or competion. As each agent comes with its own environment of granular computing <**X**, G, A>, <**Y**, G', B>, <**Z**, Q, C>. 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 3.

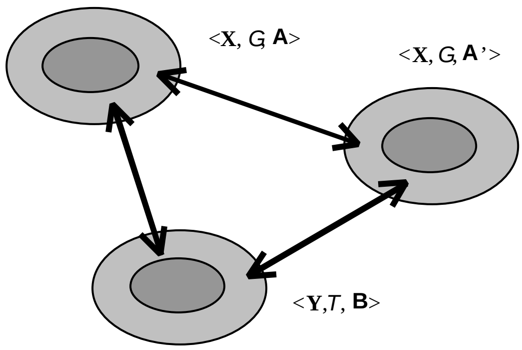

Figure 3. Collaboration between granular worlds; the mechanisms of interaction are displayed as an auxilliary processing layer around the agents.

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,

$$G = < \mathbf{X}, G, A, C > \quad (2)$$

where we mean that C may consists of a variety of constructs that help communicate (encode and decode information granules). (Bargiela, Pedrycz, 2002)

<u>*Conclusions*</u>

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 research attempted to create 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 while maintaining strong emphasis on human-centred computer applications development.

Research Impact

The work on systematisation of the Granular Computing field has already produced noticeable effects in the community. The subject of Granular Computing is addressed at most international conferences on Computational Intelligence and the number of journal papers dealing with theoretical and application aspects of Granular Computing increased significantly over the last three years. It is worth emphasising that the theoretical advances are matched by practical applications and our specific contribution to intelligent processing of travel enquiries (see below, ATTAIN - INFOHUB Ltd) makes this point.

Explanation of Expenditure

The expenditure of funds followed exactly the plan outlined in the proposal. Prof. Bargiela worked at the University of Alberta during May 2001, 2002 and 2003 and Prof. Pedrycz visited the Nottingham Trent University in December 2000, June 2001, July 2002 and August 2003.

Further Research and Dissemination Activities

Various aspects of the research completed during the project were disseminated in 8 journal papers, 21 international conference papers, 1 book (research monograph), 4 book chapters, 1 edited volume of an international journal, 1 plenary lecture, 1 keynote lecture, 1 exhibition and 5 invited lectures (3 in Japan, 1 in Italy and 1 in Finland).

Practical application of the results of Granular Computing has underpinned the formation of a spin-off company: INFOHUB Ltd. (A. Bargiela, founding director) that is focused on traffic telematics applications. (see ATTAIN at http://www.infohub-ltd.co.uk)

The practical development of the travel information system ATTAIN has been supported by grants from the Nottingham City Council and the Nottingham Traffic Control Centre. The latter is in addition to the original commitment to provide data in support of the project.

The extensive support for Prof. Pedrycz's research by the NSERC, Canada, has benefited this project by offering a very rich research ambient within which to conduct the current investigation.

The support from the Tokyo Institute of Technology, Osaka Prefecture University, Konan University, University of Bari and the Helsinki University of Technology for the invited lectures/Visiting Professorship is gratefully acknowledged.

REFERENCES

Bargiela, A., Pedrycz, W., (2002), *Granular Computing: An Introduction*, Kluwer Academic Publishers.

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., Tanaka M., (2003), A study of uncertain state estimation, *IEEE Trans. on Systems Man and Cybernetics, SMC-A*, 33, 3, 288-301.

Bargiela, A., Pedrycz, W., (2003), Recursive information granulation: Aggregation and interpretation issues, *IEEE Trans. on Systems Man and Cybernetics SMC-B*, 33, 1, 96-112

Bargiela, A., Pedrycz, W., Hirota, K. (2002), Logic-based granular prototyping, *Computers Software and Applications Conference, COMPSAC 2002*, Oxford, August 2002.

Bargiela, A. Pedrycz, W., (2002), *Archives of Control Sciences – Special Issue on Granular Computing,* ISSN 0004-072X.

Bargiela A., (2002), Granular art (interface between information sciences and arts), *Exhibition in Gallery EXTRAVAGANCE*, Sosnowiec, Poland.

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.

Pedrycz W, Bargiela A, (2003), Fuzzy fractal dimensions and fuzzy modeling, *Information Sciences*, 153, 199-216

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.