

Granularity of Knowledge, Indiscernibility and Rough Sets

Zdzisław Pawlak

Abstract— Granularity of knowledge attracted attention of many researchers recently. This paper concerns this issue from the rough set perspective. Granularity is inherently connected with foundation of rough set theory. The concept of the rough set hinges on classification of objects of interest into similarity classes, which form elementary building blocks (atoms, granules) of knowledge. These granules are employed to define basic concepts of the theory. In the paper basic concepts of rough set theory will be defined and their granular structure will be pointed out. Next the consequences of granularity of knowledge for reasoning about imprecise concepts will be discussed.

Keywords— Granularity, knowledge, rough sets, vagueness, imprecision.

I. INTRODUCTION

Information (knowledge) granulation, discussed recently by Prof. Zadeh [22, 23, 24] seems to be a very important issue for computing science, logic, philosophy and others.

In this note we are going to discuss some problems connected with granularity of knowledge in the context of rough sets (cf. Pawlak [11]). First, discussion of granulation of knowledge in connection with rough and fuzzy sets has been presented by Dubois and Prade in [5]. Recently, an interesting study of information granulation in the framework of rough sets can be found in Polkowski [12] and Skowron and Stepaniuk [16].

In rough set theory we assume that with every object some information is associated, and objects can be "seen" through the accessible information only. Hence, object with the same information cannot be discerned and appear as the same. This results in, that indiscernible objects of the universe form clusters of indistinguishable objects (granules, atoms, etc.). Thus from the rough set view the granularity of knowledge is due to the indiscernibility of objects caused by lack of sufficient information about them. Consequently granularity and indiscernibility are strictly connected and the concept of indiscernibility seems to be prior to granularity.

Indiscernibility attracted attention of philosophers for a long time and its first formulation can be attributed to Leibniz (cf. Forrest [6]), and is known as the principle of "the identity of indiscernibles". The principle says that not two objects have exactly the same properties, or in other words if all properties of objects x and y are the same then x and y are identical.

But what are "properties of objects"? and what does it mean "all properties"? A lot of philosophical discussions have been devoted to answer these questions (cf. e.g., Black

[1], Forrest [6]), but we will refrain here from philosophical debate. Let us observe only that Leibniz approach to indiscernibility identifies indiscernibility with identity. The later is obviously an equivalence relation, i.e., it leads to partition of the universe into equivalence classes (granules) of objects which are indistinguishable in view of the assumed properties. Thus in the rough set approach granulation is a consequence of the Leibniz principle.

It is worthwhile to mention that indiscernibility can be also viewed in a wider context, as pointed out by Williamson [20]: "Intelligent life requires the ability discriminate, but not with unlimited precision". This is a very interesting issue however it lays outside the scope of this paper.

In rough set theory we assume empiristic approach, i.e. we suppose that properties are simply empirical data which can be obtained as a result of measurements, observations, computations, etc. and are expressed by values of a fixed, finite set of attributes, e.g., properties are attribute-value pairs, like (size, small), (color, red) etc. The idea could be also expressed in more general terms assuming as a starting point not a set of specific attributes but abstract equivalence relation, however, the assumed approach seems more intuitive.

Equivalence relation is the simplest formalization of the indiscernibility relation and is sufficient for many applications. However, more interesting seems to assume that the indiscernibility relation is formalized as a tolerance relation, i.e., transitivity of indiscernibility is denied in this case, for, if x is indiscernible from y and y is indiscernible from z then not necessarily x is indiscernible from z . Many authors have proposed tolerance relation as a basis for rough set theory (cf. e.g., Skowron and Stepaniuk [17]). This causes, however, some mathematical complications as well philosophical questions, because it leads to vague granules, i.e., granules without sharp boundaries, closely related to the boundary-line approach to vagueness (cf. e.g., Chattebrjee [4], Sorensen [18]).

Besides, instead of tolerance relation also more sophisticated mathematical models of indiscernibility, as a basis for rough set theory, have been proposed (cf. e.g., Krawiec, Slowinski, and Vanderpooten [8], Yao and Wong, [21], Ziarko [25]). Interested readers are advised to consult the mentioned above references, but for the sake of simplicity we will adhere in this paper to the equivalence relation as a mathematical formalization of the indiscernibility relation.

Since granules of knowledge can be considered as a basic building blocks of knowledge about the universe it seems that natural mathematical model for granulated knowledge can be based on ideas similar to that used in mereology

Z. Pawlak is with Institute of Theoretical and Applied Informatics Polish Academy of Sciences ul. Baltycka 5, 44 000 Gliwice, Poland e-mail: zpw@ii.pw.edu.pl

proposed by Leśniewski [9], in which *part of* is the basic relation of this theory. Axioms of mereology, in particular in a version proposed by Suppes [19], seem to be natural candidate for this purpose. Moreover, rough mereology, extension of classical mereology proposed by Polkowski and Skowron in [13, 14], seems to be exceptionally suited to analyze granules of knowledge with not sharp boundaries (cf. Polkowski [12], Skowron and Stepaniuk [16]).

It also worthwhile to mention in this context that granularity of knowledge has been also pursued in quantum physics. Its relation to fuzzy sets and rough sets has been first mentioned by Cattaneo [2, 3].

Besides, it is also interesting to observe that computations and measurements are very good examples of granularity of information, for they are based in fact not on real numbers but on intervals, determined by the accuracy of computation or measurement.

II. BASIC PHILOSOPHY OF ROUGH SETS

The rough set philosophy is founded on the assumption that with every object of the *universe of discourse* we associate some information (data, knowledge). E.g., if objects are patients suffering from a certain disease, symptoms of the disease form information about patients. Objects characterized by the same information are *indiscernible (similar)* in view of the available information about them. The *indiscernibility relation* generated in this way is the mathematical basis of rough set theory.

Any set of all indiscernible (similar) objects is called an *elementary concepts*, and forms a basic *granule (atom)* of knowledge about the universe. Any union of some elementary concepts is referred to as *crisp (precise) concept* – otherwise the set is rough (*imprecise, vague*).

Consequently each rough concept has boundary-line cases, i.e., objects which cannot be with certainty classified neither as members of the concept nor of its complement. Obviously crisp concepts have no boundary-line elements at all. That means that boundary-line cases cannot be properly classified by employing the available knowledge.

Thus, the assumption that objects can be "seen" only through the information available about them leads to the view that knowledge has granular structure. As a consequence vague concepts, in contrast to precise concepts, cannot be characterized in terms of elementary concepts. Therefore in the proposed approach we assume that any vague concept is replaced by a pair of precise concepts – called the lower and the upper approximation of the vague concept. The lower approximation consists of all elementary concepts which *surely* are included in the concept and the upper approximation contains all elementary concepts which *possibly* are included in the concept. Obviously, the difference between the upper and the lower approximation constitutes the boundary region of the vague concept. Approximations are two basic operations in rough set theory.

III. INDISCERNIBILITY AND GRANULARITY

As mentioned in the introduction, the starting point of rough set theory is the indiscernibility relation, generated

by information about objects of interest. The indiscernibility relation is intended to express the fact that due to the lack of knowledge we are unable to discern some objects employing the available information. That means that, in general, we are unable to deal with single objects but we have to consider clusters of indiscernible objects, as fundamental concepts of knowledge.

Now we present above considerations more formally.

Suppose we are given two finite, non-empty sets U and A , where U is the *universe*, and A – a set *attributes*. With every attribute $a \in A$ we associate a set V_a , of its *values*, called the *domain* of a . The pair $S = (U, A)$ will be called an *information system*. Any subset B of A determines a binary relation I_B on U , which will be called an *indiscernibility relation*, and is defined as follows:

xI_By if and only if $a(x) = a(y)$ for every $a \in A$, where $a(x)$ denotes the value of attribute a for element x .

Obviously I_B is an equivalence relation. The family of all equivalence classes of I_B , i.e., the partition determined by B , will be denoted by U/I_B , or simply U/B ; an equivalence class of I_B , i.e., the block of the partition U/B , containing x will be denoted by $B(x)$.

If (x, y) belongs to I_B we will say that x and y are *B-indiscernible*. Equivalence classes of the relation I_B (or blocks of the partition U/B) are referred to as *B-elementary concepts* or *B-granules*.

In the rough set approach the elementary concepts are the basic building blocks (concepts) of our knowledge about reality.

IV. APPROXIMATIONS AND GRANULARITY

Now the indiscernibility relation will be used to define basic operations in rough set theory, which are defined as follows:

$$B_*(X) = \bigcup_{x \in U} \{B(x) : B(x) \subseteq X\},$$

$$B^*(X) = \bigcup_{x \in U} \{B(x) : B(x) \cap X \neq \emptyset\},$$

assigning to every $X \subseteq U$ two sets $B_*(X)$ and $B^*(X)$ called the *B-lower* and the *B-upper approximation* of X , respectively.

Hence, the *B-lower* approximation of a concept is the union of all *B-granules* that are included in the concept, whereas the *B-upper* approximation of a concept is the union of all *B-granules* that have a nonempty intersection with the concept. The set

$$BN_B(X) = B^*(X) - B_*(X)$$

will be referred to as the *B-boundary region* of X .

If the boundary region of X is the empty set, i.e., $BN_B(X) = \emptyset$, then X is *crisp (exact)* with respect to B ; in the opposite case, i.e., if $BN_B(X) \neq \emptyset$, X is referred to as *rough (inexact)* with respect to B .

Rough sets can be also defined using a *rough membership function*, defined as

$$\mu_X^B(x) = \frac{\text{card}(B(x) \cap X)}{\text{card}(B(x))}.$$

Obviously

$$\mu_X^B(x) \in [0, 1].$$

Value of the membership function $\mu_X^B(x)$ is kind of conditional probability, and can be interpreted as a degree of *certainty* to which x belongs to X (or $1 - \mu_X^B(x)$, as a degree of *uncertainty*).

The rough membership function, can be used to define approximations and the boundary region of a set, as shown below:

$$B_*(X) = \{x \in U : \mu_X^B(x) = 1\},$$

$$B^*(X) = \{x \in U : \mu_X^B(x) > 0\},$$

$$BN_B(X) = \{x \in U : 0 < \mu_X^B(x) < 1\}.$$

The rough membership function can be generalized as follows (cf. Polkowski and Skowron [13]):

$$\mu(X, Y) = \frac{\text{card}(X \cap Y)}{\text{card} X}$$

where $X, Y \subseteq U$, $X \neq \emptyset$ and $\mu(\emptyset, Y) = 1$.

Function $\mu(X, Y)$ is an example of a *rough inclusion* [14] and expresses the degree to which X is included in Y . Obviously, if $\mu(X, Y) = 1$, then $X \subseteq Y$.

If X is included in a degree k we will write $X \subseteq_k Y$.

The rough inclusion function can be interpreted as a generalization of the mereological relation "part of", and reads as "part in a degree".

Employing now the rough inclusion function we can represent approximations in an uniform way:

$$B_*(X) = \bigcup_{x \in U} \{B(x) : \mu(B(x), X) = 1\},$$

$$B^*(X) = \bigcup_{x \in U} \{B(x) : \mu(B(x), X) > 0\}.$$

Hence, the B -lower approximation of X consists of all B -granules included in X , whereas the B -upper approximation of X consists of all roughly included B -granules of X .

In this way approximations reveal granular structure of complex concepts. Those granularity of knowledge is inherently incorporated in the foundations of rough set theory.

V. DEPENDENCIES AND GRANULARITY

Another important issue in data analysis is discovering dependencies between attributes. Intuitively, a set of attributes D *depends totally* on a set of attributes C , denoted $C \Rightarrow D$, if all values of attributes from D are uniquely determined by values of attributes from C . In other words, D depends totally on C , if there exists a functional dependency between values of D and C .

We would need also a more general concept of dependency, called a *partial dependency* of attributes. Intuitively, the partial dependency means that only some values of D are determined by values of C .

Formally dependency can be defined in the following way. Let D and C be subsets of A .

We will say that D *depends on* C in a degree k ($0 \leq k \leq 1$), denoted $C \Rightarrow_k D$, if

$$k = \gamma(C, D) = \frac{\text{card}(POS_C(D))}{\text{card} U},$$

where

$$POS_C(D) = \bigcup_{X \in U/D} C_*(X),$$

called a *positive region* of the partition U/D with respect to C , is the set of all elements of U that can be uniquely classified to blocks of the partition U/D , by means of C .

Obviously

$$\gamma(C, D) = \sum_{X \in U/D} \frac{\text{card}(C_*(X))}{\text{card} U}.$$

If $k = 1$ we say that D *depends totally* on C , and if $k < 1$, we say that D *depends partially* (in a degree k) on C .

The coefficient k expresses the ratio of all elements of the universe, which can be properly classified to blocks of the partition U/D , employing attributes C and will be called the *degree of the dependency*.

Obviously if D depends totally on C then $I_C \subseteq I_D$. That means that the partition generated by C is finer than the partition generated by D .

Degree of dependency expresses to what extend granularity imposed by the set of attributes D can be expressed in terms of elementary concepts associated with C .

The function $\gamma(C, D)$ can be regarded as a generalization of the rough inclusion function $\mu(X, Y)$, for it expresses to what degree partition generated by C , i.e., U/C is included in the partition generated by D , i.e., U/D .

In other words, degree of dependency between C and D reveals to what degree granular structure imposed by D can be expressed in terms of granular structure associated with C .

In fact approximations and dependencies are different sides of the same coin, and exhibit a relationship between two kinds of granular structures.

VI. DECISION RULES AND GRANULARITY

With every dependency $C \Rightarrow_k D$ we can associate a set of decision rules, specifying decisions that should be taken when certain condition are satisfied.

To express this idea more precisely we need a formal language associated with any information system $S = (U, A)$. The language is defined in a standard way and we omit detailed definition here, assuming that the reader is familiar with the construction.

Given $x \in U$ and $B \subseteq A$ by $\Phi_x^B = \bigwedge_{a \in B} (a, v)$ we mean a formula such that $a(x) = v$ and $v \in V_a$.

Every dependency $C \Rightarrow_k D$ determines a set of *decision rules*

$$\{\Phi_x^C \rightarrow \Phi_x^D\}_{x \in U}.$$

We say that a decision rule $\Phi_x^C \rightarrow \Phi_x^D$ is *true* in S , if $|\Phi_x^C|_S \subseteq |\Phi_x^D|_S$, where $|\Phi_x^C|_S$ denotes the *meaning* of Φ_x^C in S , defined in a usual way.

Let $C_S(x) = |\Phi_x^C|_S$. Hence the decision rule $\Phi_x^C \rightarrow \Phi_x^D$ is true in S if $C_S(x) \subseteq D_S(x)$.

A decision rule $\Phi_x^C \rightarrow \Phi_x^D$ is *true in a degree l* in S , if $l = \mu(C_S(x), D_S(x)) > 0$, i.e., $C_S(x) \subseteq_l D_S(x)$.

True decision rules are often referred to as *sure* decision rules, whereas decision rules true to a degree l are called *possible* decision rules.

Let us notice that

$$\mu(C_S(x), D_S(x)) = \mu_x^C(D_S(x)),$$

Hence rough inclusion in this case boils down to rough membership function. As a consequence rough membership can be interpreted as a generalized truth value (truth to a certain degree).

The degree of truth of a decision rule can be also interpreted as a certainty factor of the rule, expressing to what degree decision imposed by the decision rule can be trusted.

It is also interesting to observe in this context that partial truth in the rough set setting has a probabilistic flavor, for the rough membership can be interpreted both as conditional probability and at the same time as partial truth value.

The above considerations lead to a D inference rule, which will be called the *rough modus ponens* and is defined below:

$$\frac{\pi(\Phi_x^C); \mu(\Phi_x^C, \Psi_x^D)}{\pi(\Psi_x^C)},$$

where

$$\pi(\Phi_x^C) = \frac{\text{card}(|\Phi_x^C|_S)}{\text{card}U},$$

$$\mu(\Phi_x^C, \Psi_x^D) = \frac{\text{card}(|\Phi_x^C \wedge \Psi_x^D|_S)}{\text{card}|\Phi_x^C|_S}$$

and

$$\pi(\Psi_x^D) = \pi(\sim \Phi_x^C \wedge \Psi_x^D) + \pi(\Phi_x^C) \cdot \mu(\Phi_x^C, \Psi_x^D).$$

The number $\pi(\Phi_x^C)$ can be interpreted as the probability, that x has the property Φ_x^C , and the number $\mu(\Phi_x^C, \Psi_x^D)$ – as *certainty factor* of the decision rule $\Phi_x^C \rightarrow \Psi_x^D$.

Hence the inference rule, the *rough modus ponens*, enables us to calculate the probability of conclusion Ψ_x^D as a simple function of the probability of the premise Φ_x^C and the certainty factor $\mu(\Phi_x^C, \Psi_x^D)$ of the decision rule $\Phi_x^C \rightarrow \Psi_x^D$.

Another approach to this problem has been proposed by Skowron in [15].

VII. CONCLUSIONS

Granularity of knowledge, information, measurements, computations etc. seems to be an intrinsic feature of our thinking and can be considered as a manifestation of an old antinomy associated with *continuos-discrete* paradigm.

Rough set philosophy hinges on the granularity of data, which is used to build all its basic concepts, like approximations, dependencies, reduction etc. Particularly interesting in this approach seems to be the relationship between partial truth and certainty, which is a natural consequence of data granularity.

ACKNOWLEDGMENTS

Thanks are due to Professor Andrzej Skowron for his critical remarks and helpful comments.

REFERENCES

- [1] M. Black, The Identity of Indiscernibles, *Mind*, 61, 1952.
- [2] G. Cattaneo, Fuzzy Quantum Logic: The Logic of unsharp Quantum Mechanics, *Int. Journal of Theoretical Physics*, 32, 1993, 1709–1734.
- [3] G. Cattaneo, Mathematical Foundations of Roughness and Fuzziness, in: S. Tsumoto et al, (eds), The fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery, Proceedings, The University of Tokyo, 1996, 241–247.
- [4] A. Chatterjee, Understanding Vagueness, Pragati Publications, Delhi, 1994.
- [5] D. Dubois, H. Prade, Foreword in: Z. Pawlak, *Rough Sets - Theoretical Aspect of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1999.
- [6] P. Forrest, Identity of Indiscernibilities, *Stanford Encyclopedia of Philosophy*, 1997.
- [7] S. French, Quantum Physics and the Identity of Indiscernibles, *British Journal of the Philosophy of Sciences*, 39, 1988.
- [8] K. Krawiec, R. Słowiński, D. Vanderpooten, Construction of Rough Classifiers Based on Application of Similarity Relation, in: Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery, November 6-8, 1996 Tokyo Japan, 23–30.
- [9] S. Leśniewski, Foundations of the General Theory of Sets, in: Surma, Srzednicki, Barnett, Riskkey (eds), Stanislaw Leśniewski Collected Works, Kluwer Academic Publishers, Dordrecht, Boston, London, 1992, 128–173.
- [10] A.F. Parker-Rhodes, The Theory of Indistinguishables, D. Reidel Publishing Company, Dordrecht, Boston, London, 1981.
- [11] Z. Pawlak, *Rough Sets - Theoretical Aspect of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1991.
- [12] L. Polkowski, A. Skowron, Towards Adaptive Calculus of Granules, manuscript, 1997.
- [13] L. Polkowski, A. Skowron, Rough Mereology, in: Proc. of the Symposium on Methodologies for Intelligent Systems, Charlotte, N.C., Lecture Notes in Artificial Intelligence, Vol. 869, Springer Verlag, Berlin, 1994, 85–94.
- [14] L. Polkowski, A. Skowron, Rough Mereology: A New Paradigm for Approximate Reasoning, *Journ. of Approximate Reasoning*, 1996, 15(4), 333–365.
- [15] A. Skowron, Management of Uncertainty in AI: A Rough Set Approach, in: V. S. Alagar, S. Bergler and F. Q. Dong, (eds.), *Incompleteness and Uncertainty in Information Systems*, Proceedings of the SOFTEHES Workshop, Concordia University, Montreal, Canada, October 8-9, 1993, 69–85.
- [16] Skowron, A., and Stepaniuk, J., Information Granulation – a Rough Set Approach, manuscript, 1997.
- [17] A. Skowron, J. Stepaniuk, Tolerance Approximation Spaces, *Fundamenta Informaticae*, 27, 1996, 245–253.
- [18] R. Sorensen, Vagueness, *Stanford Encyclopedia of Philosophy*, 1997.
- [19] P. Suppes, Some Open Problems in the Philosophy of Space and Time, *Synthese*, 24, 1972, 298–316.
- [20] T. Williamson, Identity and Discrimination, 1990, Blackwell.
- [21] Y.Y. Yao, S.K.M. Wong, Generalization of Rough Sets using Relationships between Attribute Values, in: Proceedings of the Second Annual Joint Conference on Information Sciences, Wrightsville Beach, N.C. USA, September 28 - October 1, 1995, 245–253.
- [22] L. Zadeh, Fuzzy Graphs, Rough Sets and Information Granularity, in: Proc. Third Int. Workshop on Rough Sets and Soft Computing, Nov. 10-12, 1994, San Jose.

- [23] L. Zadeh, The Key Rules of Information Granulation and Fuzzy Logic in Human Reasoning, Concept Formulation and Computing with Words, in: Proc. FUZZ-96: Fifth IEEE International Conference on Fuzzy Systems, September 8-11, 1996, New Orleans.
- [24] L. Zadeh, Information Granulation, Fuzzy Logic and Rough Sets, in: Proc. of the Fourth Int. Workshop on Rough Sets, and Machine Discovery, November 6-8, 1996, Tokyo.
- [25] W. Ziarko, Variable Precison Rough Set Model, Journal of Computer and System Sciences, 46/1, 1993, 39-59.