

VAGUENES AND UNCERTAINTY: A ROUGH SET PERSPECTIVE

ZDZISLAW PAWLAK

Institute of Computer Science, Warsaw Technical University, ul. Nowowiejska 15/19, 00 665 Warsaw, Poland and Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Baltycka 5, 44 000 Gliwice, Poland

Vagueness and uncertainty have attracted attention of philosophers and logicians for many years. Recently AI researchers contributed essentially to this area of research. Fuzzy set theory and the theory of evidence are seemingly the most appealing ones. In this note we present a new outlook on these problems, based on the rough set theory - which seems to be a very suitable mathematical tool to deal with these problems. The paper is a modified version of the authors talk entitled "An inquiry into vagueness and uncertainty", which was delivered on the AI Conference in Wigry (Poland), 1995.

Key words: vagueness, uncertainty, fuzzy sets, evidence theory, rough sets.

INTRODUCTION

The problems of vagueness and uncertainty have been for a long time considered by philosophers and logicians e.g. (Russell 1923; Russell 1950; Black 1937; Black 1963; Fine 1975). Also computer scientists, in particular, researchers interested in AI have brought new ideas to this area of research. The most important ones are seemingly the fuzzy set theory Zadeh (1965) and the theory of evidence Shafer (1976).

In this paper we would like to outline briefly another approach to these problems offered by the rough set theory Pawlak (1991). Although the proposed approach is somehow related to that offered by the fuzzy set theory Pawlak (1994) and the evidence theory, Skowron (1994) it can be viewed in its own rights.

The rough set theory bears on the assumption that we have initially some information (knowledge) about elements of the universe we are interested in. Evidently to some elements of the universe the same information can be associated and consequently the elements can be *similar or indiscernible* in view of the available information. Similarity is assumed to be a reflexive and symmetric relation, whereas the indiscernibility relation is also transitive. Thus similarity is a tolerance relation and indiscernibility is an equivalence relation.

It is worthwhile to mention in this context that the concepts of similarity and indiscernibility attracted attention of philosophers and logicians for many years e.g. Williamson (1990), nevertheless these concepts are still not understood fully. Interesting study of these problems can be also found in two recent papers of Marcus (Marcus 1994a; Marcus 1994b).

We will refrain in this paper from philosophical discussions and simply give the definitions and properties necessary to explain the ideas of vagueness and uncertainty from the rough sets perspective.

1. VAGUENES AND THE BOUNDARY REGION

Let us start from the basic definitions supporting the rough set theory.

Suppose we are given a finite not empty set U called the *universe*, and let I be a binary relation on U . By $I(x)$ we mean the set of all $y \in U$ such that yIx . If I is reflex-

ive and symmetric, i.e. xIx , for every $x \in U$ and xIy , implies yIx for every $x, y \in U$, then I is a tolerance relation. If I is also transitive, i.e. xIy and yIz implies xIz , then I is an equivalence relation. In this case $I(x) = [x]_I$, i.e. $I(x)$ is an equivalence class of the relation I containing element x . If I is a tolerance relation and xIy , then x, y are called *similar* with respects to I (*I-similar*), whereas if I is an equivalence relation and xIy , then x, y are referred to as *indiscernible* with respect to I (*I-indiscernible*). Sometimes indiscernibility is identified with similarity Marcus (1994b).

Let us define now two operations on sets

$$I_*(X) = \{x \in U : I(x) \subseteq X\},$$

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

assigning to every subset X of the universe U two sets $I_*(X)$ and $I^*(X)$ called the *I-lower* and the *I-upper approximation* of X respectively. The set

$$BN_I(X) = I^*(X) - I_*(X)$$

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

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

Thus rough sets can be viewed as a mathematical model of vague concepts.

Vagueness can be also characterized numerically by defining the following coefficient

$$\alpha_I(X) = \frac{|I_*(X)|}{|I^*(X)|},$$

where $|X|$ denotes the cardinality of the set X .

Obviously $0 \leq \alpha_I(X) \leq 1$. If $\alpha_I(X) = 1$, the set X is crisp with respect to I (the concept X is precise with respect to I), and otherwise, if $\alpha_I(X) < 1$, the set X is rough with respect to I (the concept X is vague with respect to I).

The idea of vagueness presented here is close to that expressed by Frege who fathered the so called "boundary-line" view on vagueness. According to his view

"The concept must have a sharp boundary. To the concept without a sharp boundary there would correspond an area that had not a sharp boundary-line all around". Frege (1903).

Thus Frege's idea of vagueness is due to the boundary-line cases, i.e. a vague concept must have boundary-line examples which cannot be classified, on the basis of available information, neither to the concept nor to its complement. For example the concept of an *odd (even) number* is precise, because every number is either odd or even - whereas the concept of a *beautiful women* is vague, because for some women we cannot decide whether they are beautiful or not (there are boundary-line cases).

2. UNCERTAINTY AND THE MEMBERSHIP FUNCTION

A vague concept has a boundary-line cases, i.e. elements of the universe which cannot be with *certainty* classified as elements of the concept. Hence *uncertainty*

is related to the question of *membership* of elements to a set. Therefore in order to discuss the problem of uncertainty from the rough set perspective we have to define the membership function related to the rough set concept (rough membership function), and investigate its properties.

The rough membership function can be easily defined employing the relation I in the following way:

$$\mu_X^I(x) = \frac{|X \cap I(x)|}{|I(x)|}.$$

Obviously

$$\mu_x^I(x) \in [0, 1].$$

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

$$\begin{aligned} I_*(X) &= \{x \in U : \mu_X^I(x) = 1\}, \\ I^*(X) &= \{x \in U : \mu_X^I(x) > 0\}, \\ BN_I(X) &= \{x \in U : 0 < \mu_X^I(x) < 1\}. \end{aligned}$$

Thus there exists a strict connection between vagueness and uncertainty. As we mentioned above vagueness is related to sets (concepts), whereas uncertainty is related to elements of sets, and the rough set approach shows clear connection between the two concepts, namely vagueness is defined in terms of uncertainty.

It can be shown Pawlak (1994) that the rough membership function has the following properties:

- a) $\mu_X^I(x) = 1$ iff $x \in I_*(X)$,
- b) $\mu_x^I(x) = 0$ iff $x \in U - I^*(X)$,
- c) $0 < \mu_X^I(x) < 1$ iff $x \in BN_I(X)$,
- d) If $I = \{(x, x) : x \in U\}$, then $\mu_X^I(x)$ is the characteristic function of X ,
- e) If xIy , then $\mu_X^I(x) = \mu_X^I(y)$ provided I is an equivalence relation,
- f) $\mu_{U-X}^I(x) = 1 - \mu_X^I(x)$ for any $x \in U$,
- g) $\mu_{X \cup Y}^I(x) \geq \max(\mu_X^I(x), \mu_Y^I(x))$ for any $x \in U$,
- h) $\mu_{X \cap Y}^I(x) \leq \min(\mu_X^I(x), \mu_Y^I(x))$ for any $x \in U$,
- i) If \mathbf{X} is a family of pair wise disjoint sets of U , then $\mu_{\bigcup \mathbf{X}}^I(x) = \sum_{X \in \mathbf{X}} \mu_X^I(x)$ for any $x \in U$, provided that I is an equivalence relation.

The above ideas can be generalized in the same way as proposed by Ziarko in the variable precision rough model Ziarko (1993). Let β be a real number such that $0 \leq \beta < 0.5$. Approximations can be defined now as follows:

$$\begin{aligned} I_{*\beta}(X) &= \{x \in U : \mu_X^I(x) \geq 1 - \beta\}, \\ I_{\beta}^*(X) &= \{x \in U : \mu(x) > \beta\}. \end{aligned}$$

Note that if $\beta = 0$, we get the previous case.

Remark. We could also assume that $0.5 < \beta \leq 1$ and consequently

$$I_{*\beta}(X) = \{x \in U : \mu_X^I(x) \geq \beta\},$$

$$I_{\beta}^*(X) = \{x \in U : \mu_X^I(x) > 1 - \beta\}.$$

□

The boundary region $BN_I^{\beta}(X) = I_{\beta}^*(X) - I_{*\beta}(X) = \{x \in U : \beta < \mu_X^I(x) < 1 - \beta\}$.
The following properties are obvious

- a) if $x \in I_{*\beta}(X)$, then not necessarily $I(x) \subseteq X$,
- b) if $x \in U - I_{\beta}^*(X)$, then not necessarily $X \cap I(x) = \emptyset$,
- c) if $X \cap I(x) \neq \emptyset$, then not necessarily $x \in I_{\beta}^*(X)$,
- d) if $I(x) \not\subseteq X$, then not necessarily $x \notin I_{\beta}^*(X)$.

Besides we have

$$I_*(X) \subseteq I_{*\beta}(X) \quad \text{and} \quad I^*(X) \supseteq I_{\beta}^*(X).$$

The idea of variable precision rough sets enables one to relax strict conditions superimposed on the definitions of approximations and can be viewed as a weaker form of the original definitions.

3. CLASSIFICATION OF VAGUE CONCEPTS

It turns out that the above considerations give rise to the following basic four classes of rough sets:

- a) X is *roughly I-observable*, if $0, 1 \in \{\mu_X^I(x) : x \in U\}$,
- b) X is *internally I-unobservable*, if $1 \notin \{\mu_X^I(x) : x \in U\}$ and $0 \in \{\mu_X^I(x) : x \in U\}$,
- c) X is *externally I-unobservable*, if $0 \notin \{\mu_X^I(x) : x \in U\}$ and $1 \in \{\mu_X^I(x) : x \in U\}$,
- d) X is *totally I-unobservable*, if $0, 1 \notin \{\mu_X^I(x) : x \in U\}$.

Directly from the above definitions we get the following properties:

- a) there exists $x, y \in U$ such that $I(x) \subseteq X$ and $I(y) \cap X = \emptyset$ iff X is roughly I -observable,
- b) $I(x) \not\subseteq X$ for any $x \in U$ and $I(y) \cap X = \emptyset$ for some $y \in U$, iff X is internally I -unobservable,
- c) $I(x) \cap X \neq \emptyset$ for any $x \in U$ and $I(y) \subseteq X$ for some $y \in U$, iff X is externally I -unobservable,
- d) $I(x) \not\subseteq X$ and $I(x) \cap X \neq \emptyset$ for any $x \in U$, iff X is totally I -unobservable.

The above considered sets can be also defined using the concept of approximations, as shown below:

- a) $I_*(X) \neq \emptyset$ and $I^*(X) \neq U$, iff X is *roughly I-observable*,
- b) $I_*(X) = \emptyset$ and $I^*(X) \neq U$, iff X is *internally I-unobservable*,
- c) $I_*(X) \neq \emptyset$ and $I^*(X) = U$, iff X is *externally I-unobservable*,
- d) $I_*(X) = \emptyset$ and $I^*(X) = U$, iff X is *totally I-unobservable*.

The intuitive meaning of this classification is the following:

If set X is roughly I -observable, this means that we are able to decide for some elements of U whether they belong to X or $-X$.

If X is internally I -unobservable, this means that we are able to decide whether some elements of U belong to $-X$, but we are unable to decide for any element of U , whether it belongs to X or not.

If X is externally I -unobservable, this means that we are able to decide for some elements of U whether they belong to X , but we are unable to decide, for any element of U whether it belongs to $-X$ or not.

If X is totally I -unobservable, we are unable to decide for any element of U whether it belongs to X or $-X$.

That means, that the set X is roughly observable if there are some elements in the universe which can be positively classified, to the set X . This definition also implies that there are some other elements which can be classified without any ambiguity as being outside the set X .

External I -unobservability of a set refers to the situation when positive classification is possible for some elements, but it is impossible to determine that an element does not belong to X .

The above classification leads to the conclusion that there are four natural classes of vagueness and corresponding uncertainties. It turns out that besides theoretical importance the above classification has also great practical significance when using the rough set theory to data analysis Slowinski (1992).

CONCLUSION

The rough set theory seems to be a natural mathematical model of vagueness and uncertainty. Vagueness is a property of sets (concepts) and can be attributed to the boundary region of a set, whereas uncertainty is a property of elements of sets and is related to the rough membership function.

The presented approach can be seen as complementary view of the problems considered to that offered by the fuzzy set theory and the evidence theory.

ACKNOWLEDGEMENTS

Thanks are due to Professor Andrzej Skowron for critical remarks.

REFERENCES

- Black, M. 1937. Vagueness. *The Philosophy of Sciences*, **2**, pp. 427-455.
 Black, M. (1963). Reasoning with Loose Concepts. *Dialog*, **2**, pp. 1-12.
 Fine, K. (1975). Vagueness, Truth and Logic. *Synthese*, **30**, pp. 265-300.
 Frege, G. (1903). *Grundgesetze der Arithmentik*, **2**, in Geach and Black Eds.) *Selections from the Philosophical Writings of Gotlob Frege*, Blackwell, Oxford 1970.
 Hemple, C., G. (1952). *Fundamental of Concept Formation in Empirical Sciences*. University of Chicago Press, Chicago.
 Hunt, E., B. (1974). *Concept Formation*. John Wiley and Sons, New York.
 Marcus, S. (1994). *it Imprecision between Variety and Uniformity: The Conjugate Pairs*, ICS Research Reports 50/93 and Bull. Pol. Acad. Sci. (to appear).

- Marcus, S. (1994). Tolerance Rough Sets, Āech Topology, Learning Processes, ICS Research Reports 51/93 and Bull. Pol. Acad. Sci. (to appear).
- Pawlak, Z., Wong S.,K.,M., and Ziarko, W. (1988). it Rough Sets: Probabilistic Versus Deterministic Approach. Int. J. Man-Machine Studies, **29**, pp. 81-95
- Pawlak, Z. (1991). it Rough Sets - Theoretical Aspects of Reasoning about Data. KLUWER ACADEMIC PUBLISHERS.
- Pawlak, Z., and Skowron, A. (1994). Rough Membership Functions. In:Yager, R.R., Fedrizzi, M., and Kacprzyk, J. (Eds.), Advances in the Dempster Shafer Theory of Evidence, John Wiley and Sons, pp. 251-271.
- Popper, K. (1959). The Logic of Scientific Discovery. London: Hutchinson.
- Russell, B. (1923). Vagueness. Australian Journal of Philosophy, **1**, pp. 84-92.
- Russell, B.(1950). An Inquiry into Meaning and Truth. George Allen and Unwin, London.
- Shafer, G. (1976). A Mathematical Theory of Evidence. Princeton, NJ: Princeton University Press.
- Skowron, A., and Rauszer, C. (1992). The Discernibility Matrices and Functions in Information Systems. In : Slowinski, R. (Ed.) Intelligent Decision Support - Handbook of Advances and Applications of the Rough Set Theory. KLUWER ACADEMIC PUBLISHERS, pp. 311-362.
- Skowron, A., and Grzymala-Busse, J. (1994). From the Rough Set Theory to Evidence Theory. In: Yaeger, R.R., Fedrizzi, M., and Kacprzyk, J. (Eds.), Advances in the Dempster Shafer Theory of Evidence, John Wiley and Sons, pp. 193-235.
- Skowron, A., and Stepaniuk, J. (1994). Approximation of Relations (to appear).
- Slowinski, R. (Ed.) (1992). Intelligent Decision Support - Handbook of Advances and Applications of the Rough Set Theory. KLUWER ACADEMIC PUBLISHERS.
- Williamson, T. (1990). Identity and Discrimination .Basil Blackwell, Cambridge, Massachusetts.
- Zadeh, L. (1965). Fuzzy Sets. Information and Control **8**, pp.338-353.
- Ziarko, W. (1993). Variable Precision Rough Set Model. Journal of Computer and System Sciences, vol. **40(1)**, pp. 39-59.