Motto :*"Reality or the world we all know, is only a description".*
Carlos Castaneda, in Journey to Ixtlan :The lesson of Don Juan

KNOWLEDGE, REASONING AND CLASSIFICATION
A ROUGH SET PERSPECTIVE

Zdzisław Pawlak

Institute of Theoretical and Applied Informatics

Polish Academy of Sciences

Warsaw

1. Introduction.

Knowledge is widely discussed issue nowadays, mainly by logicians and computer scientists, in connection with artificial intelligence (cf. Brachman at al.(1985), Halpern (1986)). Intuitively, knowledge can be perceived as a body of information about some parts of reality, which constitute our domain of interest. This definition, however fails to meet precision standards needed when developing a formal theory. Moreover, at closer look it has multiple meanings, and it tends to mean one of several things depending on the context and the area of interest.

We propose here a formal definition of the term "knowledge" and we show some its basic properties employing the rough set concept. We realize that the proposed understanding of knowledge is not sufficiently general, yet it seems to cover variety of domains, in particular in computer science, and artificial intelligence.

The concept of knowledge presented here is rather closer to that considered in cognitive sciences, than that discussed in artificial intelligence. We do not, however, have aim to form a new, general theory of knowledge, but we have in mind practical applications and in fact the proposed approach has been successfully applied in many areas (cf. Arciszewski (1986), Krasowski (1988), Mrózek (1987), Słowiński (1988), Ziarko (1989), and others).

2. Knowledge as Classification.

Our claim is that knowledge is deep-seated in the classificatory abilities of human beings and other species. For example, knowledge about the environment is primarily manifested as an ability to classify variety of situations from the point of view of survival in the real world. Complex classification patterns of sensors signals probably form fundamental mechanisms of every living being. Classification on more abstract levels, seems to be key issue in reasoning, learning and decision making, not to mention that in science classification is of primary importance, too.

We simply assume here, that knowledge is the ability to classify objects, and by object we mean anything we can think of, for example, real things, states, abstract concepts, processes etc. In fact knowledge consists of family of various classification patterns, of a domain of interest. Thus by knowledge we mean here variety of classification patterns related to specific parts of real or abstract world, called here *the universe of discourse* (in short *the universe*). Nothing particular about the nature of the universe and knowledge will be assumed in this paper.

In what follows we shall explain this idea in some more detail.

3. Knowledge Base.

For mathematical reasons we shall often use equivalence relations instead of classifications, since these two notions are mutually exchangeable and the later is easier to deal with. Hence knowledge can be also understood as a family of equivalence relations over a fixed universe. To define the idea more precisely we need some formal definitions which are given below.

By *knowledge base* we mean relational system $K = (U, R)$, where U is a finite set called *the universe*, and R is a fa-

mily of equivalence relations over U. Every $P \subseteq R$ is called *knowledge about U*.

If ρ is an equivalence relation, then by ρ/U we mean the family of all equivalence classes of ρ, referred to as *categories of* ρ (or ρ/U).

If $\rho \in R$, then ρ will be called *primitive* and categories of ρ are also said to be *primitive*.

For example if the elements of the universe are categorized according to color, then the corresponding categories would be specific colors, for instance green, red etc.

If R is knowledge about U and $P \subseteq R$, then

$$E(P) = \bigcap_{\rho \in P} \rho$$

is also an equivalence relation over U, and

$$[x]_{E(P)} = \bigcap_{\rho \in P} [x]_\rho$$

where $[x]_\rho$ denotes category of ρ containing element x.

Categories of $E(P)$ will be also called *atoms* of P (*P-atoms*), or *elementary categories* of P (*P-elementary categories*).

Thus elementary categories of P consist of those objects of the universe U which are *similar*, or *indiscernible* according to knowledge P. In other words elementary categories of P are those elementary properties of the universe which can be voiced employing knowledge P.

Knowledge P and Q are equivalent ($P \simeq Q$), if $E(P) = E(Q)$.

Hence $P \simeq Q$, if both P and Q have the same set of elementary categories.

If $E(P) \subseteq E(Q)$ we say that knowledge P is *finer* than knowledge Q, or Q is *coarser* than P. We can also say, that if P is finer than Q, then P is *specification* of Q, and Q is *generalization* of P.

Every union of elementary categories of P will be called *category* of P (*P-category*).

Thus *P-category* is a property which can be expressed in terms of P-atoms, i.e. elementary properties expressible in knowledge P.

It is easily seen that all categories form a boolean algebra, i.e. categories are closed under boolean operations. For example if "old" and "ill" are categories, then " old and ill", " old or ill" and " not ill" are also categories in our language.

4. Categories and Concepts.

Suppose we are given a knowledge base $K = (U, R)$. Any subset $X \subseteq U$ will be called a *concept* in K. We will say that the concept X is *definable* in knowledge $P \subseteq R$ (*is P-definable*), if X is a category of P (i.e. union of elementary categories of P); otherwise the concept X is non-definable in P.

(Another, very elegant, mathematical analysis of concepts was proposed by Wille (1982). See also Iwiński (1989)).

The P-definable concepts are those which can be exactly expressed employing the knowledge P, and will be also called *exact* in P (*P-exact*). The P-nondefinable concepts cannot be expressed using the knowledge P and are said to be *inexact* or *rough* in P (*P-rough*).

Rough concepts can be however defined approximately and to this end we assign two categories which will approximate the concept from below and from above. These two categories of knowledge P are of course definable in P and will be referred to as a lower and upper approximation of X in P. Formal definition of approximations is given next.

Let $P \subseteq R$ and $X \subseteq U$. The *P-lower approximation of* X, denoted $\underline{P}X$, and the *P-upper approximation* of X, denoted $\overline{P}X$ are, defined as below.

$$\underline{P}X = \bigcup\{Y \in E(P)/U:\ Y \subseteq X\}$$

$$\overline{P}X = \bigcup\{Y \in E(P)/U:\ Y \cap X \neq \emptyset\}$$

The set $BN_P(X) = \overline{P}X - \underline{P}X$. will be called a *boundary* of X in P

Set $\underline{P}X$ is the set of all elements of U which can be with *certainty* classified as elements of X, employing knowledge P; Set $\overline{P}X$ is the set of elements of U which can be *possibly* classified as elements of X, using the knowledge P. Set $BN_P(X)$ is the set of elements which cannot be classified either to X or to $-X$ by using knowledge P.

It is easy to see that the concept X is exact in P, if and only if $\underline{P}X = \overline{P}X$; otherwise, i.e. if $\underline{P}X \neq \overline{P}X$, X is rough in P. (cf. Pawlak (1982)).

One can easily prove the following properties of approximations.

Proposition 1.

1) $\underline{P}X \subseteq X \subseteq \overline{P}X$

2) $\underline{P}\emptyset = \overline{P}\emptyset = \emptyset$; $\underline{P}U = \overline{P}U = U$

3) $X \subseteq Y$ implies $\underline{P}X \subseteq \underline{P}Y$ and $\overline{P}X \subseteq \overline{P}Y$

4) $\overline{P}(X \cup Y) = \overline{P}X \cup \overline{P}Y$

5) $\underline{P}(X \cap Y) = \underline{P}X \cap \underline{P}Y$

6) $\underline{P}(X \cup Y) \supseteq \underline{P}X \cup \underline{P}Y$

7) $\overline{P}(X \cap Y) \subseteq \overline{P}X \cap \overline{P}Y$

8) $\underline{P}(-X) = -\overline{P}X$

9) $\overline{P}(-X) = -\underline{P}X$

10) $\underline{P}\,\underline{P}X = \overline{P}\,\underline{P}X = \underline{P}X$

11) $\overline{P}\,\overline{P}X = \underline{P}\,\overline{P}X = \overline{P}X$

5. Reduction of Knowledge.

Fundamental problem we are going to address in this section is whether the whole knowledge is always necessary to define some categories available in the knowledge considered. This problem arises in many practical applications and will be referred to as knowledge reduction.

To discuss the problem precisely we define first some notions.

Let $K = (U, R)$ be a knowledge base and let $P, Q \subseteq R$.

By *P-positive region* of Q, denoted $POS_P(Q)$ we understand the set

$$POS_P(Q) = \bigcup_{X \in E(Q)/U} \underline{P}X$$

We say that $\rho \in P$ is *Q-indispensable in P*, if

$$POS_P(Q) \neq POS_{P-\{\rho\}}(Q);$$

otherwise ρ is *Q-dispensable in P*.

If every ρ in P is *Q-indispensable* we will say that P is *Q-independent*.

The set of all Q-indispensable primitive relations in P will be called the *Q-core knowledge* of P, and will be denoted as $CORE_Q(P)$.

Set $S \subseteq P$ will be called a *Q-reduct* of P, if S is Q-independent subset of P and $POS_S(Q) = POS_P(Q)$.

The following property explains the relationship between the core and reducts.

Proposition 2.

$$CORE_Q(P) = \bigcap_{R \in RED_Q(P)} R$$

where $RED_Q(P)$ is the family of all Q-reducts of P. ■

Let us briefly comment the above defined notions.

Set $POS_P(Q)$ is the set of all objects which can be classified to elementary categories of knowledge Q, employing knowledge P.

Knowledge P is Q-independent if the whole knowledge P is necessary to classify objects to elementary categories of knowledge P.

The Q-core knowledge of P is the most essential part of knowledge P, which cannot be eliminated without disturbing the ability to classify objects to elementary categories of Q.

The Q-reduct of knowledge P is minimal subset of knowledge P, which provides the same classification of objects to elementary categories of knowledge Q as the whole knowledge P. Let us observe that knowledge P can have more than one reduct.

Knowledge P with only one Q-reduct is in a sense deterministic, i.e. there is only one way of using elementary categories of knowledge P when classifying objects to

elementary categories of knowledge Q. In the case of nondeterministic knowledge i.e. if knowledge P has many Q-reducts, there are, in general, many ways of using elementary categories of P when classifying objects to elementary categories of Q. This nondeterminism is particularly strong if the core knowledge is void. Hence nondeterminism introduces synonymy to the knowledge, which in some cases may be a drawback.

6. Reasoning about Knowledge.

Theorizing, besides classification, is the second most important aspect when drawing inferences about the world. Essentially, developing theories is based on discovering inference rules of the form "if ... then ". (Sometimes the rules can describe causal relationships). In our philosophy this can be formulated as how from a given knowledge another knowledge can be induced.

More precisely, knowledge Q is *derivable* from knowledge P, if all elementary categories of Q can be defined in terms of some elementary categories of knowledge P. The derivation can be also partial, which means that only part of knowledge Q is derivable from P. The partial derivability can be defined using the notion of positive region of knowledge.

We define now the partial derivability formally.

Let $K = (U,R)$ be knowledge base and $P,Q \subseteq R$. We say that knowledge Q is *derivable in a degree* k $(0 \leq k \leq 1)$ from knowledge P, symbolically $P \longrightarrow^k Q$, if

$$k = \gamma_P(Q) = \frac{card\ POS_P(Q)}{card\ (U)}$$

where card denotes cardinality of the set.

If $k = 1$ we will say that Q *is totally derivable from* P; if $0<k<1$, we say that Q *is roughly (partially) derivable from* P, and if $k = 0$ we say that Q is *totally nonderivable from* P. If $P \longrightarrow^1 Q$ we shall also write $P \dashrightarrow Q$.

The above described ideas can be also interpreted as ability to classify objects. More precisely, if $k = 1$, then all elements of the universe can be classified to elementary categories of Q by using knowledge P. If $k \neq 1$ only those element of the universe which belong to the positive region can be classified to categories of knowledge Q, employing knowledge P. In particular, if $k = 0$ none of the elements of the universe can be classified , using knowledge P – to elementary categories of knowledge Q.

7. Knowledge Representation.

The notion of knowledge we have been considering so far has semantic character. For computational reasons we need however syntactic representation of knowledge. To this end we shall employ tabular representation of knowledge, which can be viewed as a special kind of "formal language" used to represent equivalence relations (or partitions) in symbolic form suitable for computer processing. Such a data table will be called knowledge representation systems. (Sometimes called also information systems or attribute-value system).

Knowledge representation system can be perceived as a data table, columns of which are labelled by attributes, rows are labelled by objects (states, processes etc.) and each row represents a piece of information about the corresponding object. The data table can be obtained as a result of measurements, observations or represents knowledge of an agent or a group of agents. The origin of the data table is not important from our point of view and we shall be interested only in some formal properties of such tables.

It is easily seen that with each attribute we can associate an equivalence relation. For example the attribute "color" classifies all objects of the universe into categories of objects having the same color, like red, green , blue etc. Hence with the whole table we can associate the set of equivalence relations, i.e. knowledge base. All the problems mentioned in the previous paragraphs can be now formulated in terms of classifications induced by attributes and their values.

For example knowledge reduction and reasoning about knowledge can be formulated as reduction of attributes and detecting attribute (partial) dependencies.

Formally knowledge representation system can be formulated as follows.

Knowledge representation system is a pair $S = (U, A)$, where

U – is a nonempty, finite set called the *universe*.

A – is a nonempty, finite set of *primitive attributes*.

Every primitive attribute $a \in A$ is a total function $a: U \longrightarrow V_a$, where V_a – is the set of *values* of a, called the *domain* of a.

With every subset of attributes $B \subseteq A$, we associate a binary relation $E(B)$ defined thus:

$$E(B) = \{(x, y) \in U^2 : \forall_{a \in B}\ a(x) = a(y)\}.$$

Obviously $E(B)$ is an equivalence relation and

$$E(B) = \bigcap_{a \in B} E(a)$$

Every subset $B \subseteq A$ will be called an *attribute*. If B is a single element set then B is a primitive attribute, otherwise the attribute is said to be *compound*. Attribute B may be considered as a name of the relation $E(B)$. For simplicity, if it does not cause confusion, we shall identify attributes and the corresponding relations.

The value $a(x)$ assigned by the primitive attribute a to the object x can be viewed as a name (or a description) of the primitive category of a to which x belongs (i.e. equivalence class of $E(a)$ containing x, that is to say $a(x)$ is the name of $[x]_{E(a)}$). The name (description) of an elementary category of attribute $B \subseteq A$ containing object x is a set $\{a(x)\}_{a \in B}$.

The example below will illustrate the notions defined so far.

8. Example.

Let us consider the following knowledge representation system.

U	a	b	c	d	e
1	1	0	2	2	0
2	0	1	1	1	2
3	2	0	0	1	1
4	1	1	0	2	2
5	1	0	2	0	1
6	2	2	0	1	1
7	2	1	1	1	2
8	0	1	1	0	1

Table 1

The universe U consists of 8 elements 1,2,3,4,5,6,7,8, the set of attributes is $A = \{a,b,c,d,e\}$, whereas

$$V = V_a = V_b = V_c = V_d = V_e = \{0,1,2\}.$$

It should be quite clear that some objects may have identical values of some attributes, i.e. they cannot be distinguished by these attributes, what means that they must belong to the same elementary category of the set of attributes. Exemplary partitions generated by attributes in this system are given below.

a - {{2,8},{1,4,5},{3,6,7}}

b - {{1,3,5},{2,4,7,8},{6}}

$\{c,d\}$ - {{1},{2},{3,6},{2,7},{4},{5},{8}}

$\{a,b,c\}$ - {{1,5},{2,8},{3},{4},{6},{7}}

Reduction of knowledge can be now expressed as elimination of superfluous attributes, and rules of inference can be interpreted as partial dependencies of attributes in the knowledge representation system.

Suppose that the subsets of attributes $\{c,d,e\}$ and $\{a,b\}$, in Table 1, represent knowledge P and Q respectively. Then $POS_P(Q) = \{1,4,5,8\}$, what means that only those objects can be classified to elementary categories of Q, employing knowledge P, or in other words there is a partial dependency between Q and P, or more exactly $P \rightarrow^{0.5} Q$, i.e 50% object only can be classified. Moreover knowledge P is Q-dependent; the Q-core of P is the attribute c and there are two Q-reducts of P, namely $\{c,d\}$ and $\{c,e\}$. Because knowledge P

is nondeterministic, we can use either reduct $\{c,d\}$ or $\{c,e\}$ to classify objects to categories of Q.

Tabular "language" of knowledge representation leads to simple algorithms for knowledge reduction and inference rules (dependencies) generation. More about this topics can be found in Grzymala (1988) and Ziarko (1987).

9. Conclusion.

The presented approach to reasoning about knowledge, based on the rough set philosophy, allows of simple and precise formulation of problems involving imprecision. It seems to be of particular interest for artificial intelligence, in particular, machine learning, pattern recognition, decison support systems, decision tables simplification and others, since they can be reduced to the above scheme of knowledge processing. More details can be found in specialistc literature.

References.

Arciszewski,T. & Ziarko,W, (1986). Adaptive Expert System for Preliminary Engineering Design. *Proc. 6-th International Workshop on Expert Systems and their Applications. Avignon,France,695-712.*

Brachman,R.J. & Levesque,H.J. (ed). (1986). Readings in Knowledge Representation. *Morgan Kaufmann Publishers. Inc.*

Fibak,J.Słowiński,K. & Słowiński,R. (1986). The Application of Rough Set Theory to the Verification of Indications for Treatment of Duodenal Ulcer By HSV. *Proc. 6-th International Workshop on Expert Systems and their Applications. Avignon. France, 587-599.*

Grzymala-Busse,J.W. (1988). Knowledge Acquisition under Uncertainty - a Rough Set Approach. *Journal of Intelligent and Robotic Systems. Vol.1.No.1. pp. 3-16.*

Halpern,J.Y. (ed.).(1986). Theoretical Aspects of Reasoning About Knowledge. *Proc. of the 1986 Conference. March 19-22. Monterey, California, Morgan Kaufmann Publishers. Inc.*

Hunt,E.D., Martin,J. & Stone,P.J. (1986). Experiments in Induction. *Academic Press, New York.*

Iwiński,T. (1989). Rough Orders, Rough Concepts. *Bull. Pol. Acad. Sci. Math.,(to appear).*

Krasowski,H. (1988). Aircraft Pilot Performance Evaluation with Rough Sets. *Ph.D.Dissertation.*

Mrózek, A. (1987). Rough Sets and Some Aspects of Expert Systems Realizations. *Proc. 7-th International Workshop on Expert Systems and their Applications. Avignon, France, 597-611.*

Novotny, M. & Pawlak, Z. (1989). Partial Dependency of Attributes. *Bull. Pol. Acad. Sci. Math., (to appear).*

Novotny, M. & Pawlak, Z. (1989). Independence of Attributes. *Bull. Pol. Acad. Sci. Math., (to appear).*

Orłowska, E. (1983). Dependencies of attributes in Pawlak's Information Systems. *Fundamenta Informaticae, 6, (3-4), pp. 247-256.*

Pawlak, Z. (1981). Information Systems - Theoretical Foundations. *Information Systems, 6, 205-218.*

Pawlak, Z. (1982). Rough Sets. *International Journal of Information and Computer Sciences, 11, 341-356.*

Pawlak, Z. (1985). On Rough Dependency of Attributes in Information Systems. *Bull. Pol. Acad. Sci. Tech., 33,* pp. 551-559.

Pawlak, Z., Słowiński, K. & Słowiński, R. (1986). Rough Classification of Patients after Highly Selective Vagotomy for Duodenal Ulcer. *Int. J. Man-Machine Studies. 24, 413-433.*

Rauszer, C. M. (1989). Algebraic Properties of Functional Dependencies. *Bull. Pol. Acad. Sci. Math. (to appear).*

Słowiński, K., Słowiński, R. & Stefanowski, R. (1988). Rough Sets Approach to Analysis of Data from Peritoneal Lavage in Acute Pancreatitis. *Medical Informatics. Vol. 13, No 3, pp. 143-156.*

Wille, R. (1982). Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts. *In: I. Rival (Ed.), Ordered Sets, Reidel, Dordrecht - Boston, pp. 445-470.*

Ziarko, W. (1987). On Reduction of Knowledge Representation. *Proc. 2-nd Inter. Symp. on Methodologies for Intelligent Systems. Charlotte, NC. North Holland.*

Ziarko, W. & Katzberg. J. D. (1989). Control Algorithm Acquisition, Analysis and Reduction: a Machine Learning Approach. *Knowledge Based System Diagnosis, Supervision and Control, Plenum Press, (to appear).*