A Treatise on Rough Sets

Zdzisław Pawlak

[1] Institute for Theoretical and Applied Informatics,
Polish Academy of Sciences, Bałtycka 5, 44-100 Gliwice, Poland
[2] Warsaw School of Information Technology, Newelska 6, 01-447 Warsaw, Poland
zpw@ii.pw.edu.pl

The central problem of our age is how to act decisively in the absence of certainty.
Bertrand Russell (1950). An Inquiry into Meaning and Truth.
George Allen and Unwin, London;
W.W. Norton, New York

Abstract. This article presents some general remarks on rough sets and their place in general picture of research on vagueness and uncertainty - concepts of utmost interest, for many years, for philosophers, mathematicians, logicians and recently also for computer scientists and engineers particularly those working in such areas as AI, computational intelligence, intelligent systems, cognitive science, data mining and machine learning. Thus this article is intended to present some philosophical observations rather than to consider technical details or applications of rough set theory. Therefore we also refrain from presentation of many interesting applications and some generalizations of the theory.

Keywords: Sets, fuzzy sets, rough sets, antinomies, vagueness.

1 Introduction

In this article we are going to give some general remarks on rough sets and their place in general picture of research on vagueness and uncertainty - concepts of utmost interest, for many years, for philosophers, mathematicians, logicians and recently also for computer scientists and engineers particularly those working in such areas as AI, computational intelligence, intelligent systems, cognitive science, data mining and machine learning. Thus this article is intended to present some philosophical observations rather than to consider technical details or applications of rough set theory. Therefore we also refrain from presentation of many interesting applications and some generalizations of the theory.

We start our consideration in Section 2 with general comments on classical notion of a set, formulated by Georg Cantor [8] over one hundred years ago. Next, we discuss briefly in Section 3 a source of basic discomfort of classical set theory, namely the antinomies, which shocked the foundation of mathematics

J.F. Peters and A. Skowron (Eds.): Transactions on Rough Sets IV, LNCS 3700, pp. 1–17, 2005.

and the ways out of this embarrassment. Further, the notion of vagueness and its role in mathematics, as formulated by Gottlob Frege [12] are briefly discussed in Section 4. The basic notions concerning fuzzy sets and rough sets are presented in Sections 5 and 6, respectively. The contrast between fuzzy membership [51] and rough membership [31] is briefly considered in Section 7. Than we discuss the notions of fuzzy set [51] and rough set [28,30] as certain formalizations of vagueness. A brief comparison of both notions close this section.

We conclude our deliberation in Section 8 with brief discussion of deductive, inductive and common sense reasoning and the role of rough sets has played in these kinds of inference.

2 Sets

The notion of a set is the basic one of mathematics. All mathematical structures refer to it.

The definition of this notion and the creation of set theory are due to German mathematician Georg Cantor (1845-1918), who laid the foundations of contemporary set theory about 100 years ago. The original, intuitive definition of the Cantor's notion of the set [8] is given below:

"Unter einer "Manningfaltigkeit" oder "Menge" verstehe ich nämlich allgemein jedes Viele, welches sich als Eines denken lässt, d.h. jeden Inbegriff bestimmter Elemente, welcher durch ein Gesetz zu einem Ganzen verbunden werden kann."

Thus according to Cantor a set is a collection of any objects, which according to some law can be considered as a whole. As one can see the notion is very intuitive and simple.

All mathematical objects, e.g., relations, functions, numbers, etc. are some kind of sets. In fact set theory is needed in mathematics to provide rigor.

The notion of a set is not only fundamental for the whole mathematics but it also plays an important role in natural language. We often speak about sets (collections) of various objects of interest such as, collection of books, paintings and people.

The intuitive meaning of a set according to some dictionaries is the following:

"A number of things of the same kind that belong or are used together."

Webster's Dictionary

"Number of things of the same kind, that belong together because they are similar or complementary to each other."

The Oxford English Dictionary

Thus a set is a collection of things which are somehow related to each other but the nature of this relationship is not specified in these definitions.

In fact, these definitions are due to the original definition given by Cantor.

3 Antinomies

Well! I'have seen often a cat without a grin, thought Alice; but a grin without a cat!
Lewis Carroll (1994). Alice's Adventures in Wonderland.
Penguin Books, London

In 1903 the renown English philosopher Bertrand Russell (1872-1970) observed [37] that the intuitive notion of a set given by Cantor leads to logical *antinomies* (contradictions), i.e., set theory was contradictory (there also exist other kinds of antinomies - we refrain from considering them here). A logical antinomy, for the sake of simplicity called antinomy in the remaining part of this paper, arises when after carrying on a correct logical reasoning we come to a contradiction, i.e., to the propositions A and *non-A*, which is not allowed in logic.

As an example let us discuss briefly the so-called Russell's antinomy. Consider the set X containing all the sets Y, which are not the elements of themselves. If we assume that X is its own element then X, by definition, cannot be its own element; while if we assume that X is not its own element then, according to the definition of the set X, it must be its own element. Thus while applying each assumption we obtain contradiction.

The above antinomy is often illustrated with the example of a barber, who got the instruction, that he could only shave all the men who did not shave themselves. Then a question arises if he may shave himself or not. If we assume that the barber shaves himself then, according to the instruction, he may not shave himself. But when we assume that he does not shave himself then, according to the instruction, he should shave himself. Thus we have run across an antinomy.

Another well known antinomy, called the power-set antinomy, goes as follows: consider (infinite) set X of all sets. Thus X is the greatest set. Let Y denote the set of all subsets of X. Obviously Y is greater then X, because the cardinality of the family of all subsets of a given set is always greater the cardinality of the set of all its elements. For example, if $X = \{1, 2, 3\}$ then $Y = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$, where $\emptyset$ denotes the empty set. Hence, X is not the greatest set as assumed and we have arrived at contradiction.

Antinomies show that a set cannot be a collection of arbitrary elements, as was stipulated by Cantor.

One could think that antinomies are ingenuous logical play, but it is not so. They question the essence of logical reasoning. That is why there have been attempts to "repair" Cantor's theory for over 100 years or to substitute another set theory for it but the results have not been good so far. Is then all mathematics based on doubtful foundations?

As a remedy for this defect several improvements of set theory have been proposed. For example,

- Axiomatic set theory (Zermello and Fraenkel, 1904);
- Theory of types (Whitehead and Russell, 1910);
- Theory of classes (v. Neumann, 1920).

All of these improvements consist in restrictions put on objects which can form a set. Such restrictions are expressed by properly chosen axioms, which say how a set can be built. They are called, in contrast to Cantors' intuitive set theory, axiomatic set theories.

Instead of improvements of Cantors' set theory by its axiomatization, some mathematicians proposed escape from classical set theory by creating a completely new idea of a set, which would free the theory from antinomies. Some of them are listed below.

– Mereology (Leśniewski, 1915, [19]);
– Alternative set theory (Vopenka, 1970, [49]);
– "Penumbral" set theory (Apostoli and Kanada, 1999, [1]).

No doubt the most interesting proposal was given by Polish logician Stanisław Leśniewski, who proposed instead of membership relation between elements and sets, employed in classical set theory, the relation of "being a part". In his set theory, called *mereology*, this relation is a fundamental one [19].

None of the three mentioned above "new" set theories were accepted by mathematicians. However, Leniewski's mereology attracted some attention of philosophers and recently also researchers in computer science (see, e.g., [9,33,43]).

The problem of finding an alternative to classical set theory has failed to be solved until now.

Basic concept of mathematics, the set, leads to antinomies, i.e., it is contradictory. How is it is then possible that mathematics is so successful and can be applied almost everywhere – that bridges are not collapsing, air-planes are not falling down and man has landed on the moon?

The deficiency of sets, mentioned above, has rather philosophical than practical meaning, since sets used practically in mathematics are free from the above discussed faults. Antinomies are associated with very "artificial" sets constructed in logic but not found in sets used in "everyday" mathematics. That is why we can use mathematics safely.

4 Vagueness

Besides known and unknown what else is three?
Harold Pinter (1965). The Homecoming.
Methuen, London

Another issue discussed in connection with the notion of a set is vagueness. Mathematics requires that all mathematical notions (including set) must be exact, otherwise precise reasoning would be impossible. However, philosophers [17,18,36,38] and recently computer scientists [21,23,24,41] as well as other researchers have become interested in *vague* (imprecise) concepts.

In classical set theory a set is uniquely determined by its elements. In other words, this means that every element must be uniquely classified as belonging to the set or not. That is to say the notion of a set is a *crisp* (precise) one. For example, the set of odd numbers is crisp because every number is either odd or even.

In contrast to odd numbers, the notion of a beautiful painting is vague, because we are unable to classify uniquely all paintings into two classes: beautiful and not beautiful. Some paintings cannot be decided whether they are beautiful or not and thus they remain in the doubtful area. Thus *beauty* is not a precise but a vague concept.

Almost all concepts we are using in natural language are vague. Therefore common sense reasoning based on natural language must be based on vague concepts and not on classical logic. This is why vagueness is important for philosophers and recently also for computer scientists. Interesting discussion of this issue can be found in [36].

The idea of vagueness can be traced back to ancient Greek philosophers Eubulides (ca. 400BC) who first formulated so called sorites (Bald Man or Heap) paradox. The paradox goes as follows: suppose a man has 100 000 hair on his head. Removing on hair from his head surely cannot make him bald. Repeating this step we arrive at the conclusion the a man without any hair is not bald. Similar reasoning can be applied to a hip stones.

Vagueness is usually associated with the boundary region approach (i.e., existence of objects which cannot be uniquely classified relative to a set or its complement) which was first formulated in 1893 by the father of modern logic, German logician, Gottlob Frege (1848-1925). He wrote:

"Der Begrieff muss scharf begrenzt sein. Einem unscharf begrenzten Begriff würde ein Bezirk ensprechen, der nicht überall ein scharfe Grentzlinie hätte, sondern stellenweise gantz verschwimmend in die Umgebung übergine" [12].

Thus according to Frege:

"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."

It means mathematics must use crisp, not vague concepts, otherwise it would be impossible to reason precisely. Summing up, vagueness is

– not allowed in mathematics;
– interesting for philosophy;
– necessary for computer science.

5 Fuzzy Sets

There is nothing new under the sun.
Ecclesiates 1:9

At the same time, independently of mathematicians' and philosophers' investigations, engineers became interested in the notion of a set. It turned out that many practical problems could not be formulated and solved by means of classical Cantor's notion of a set.

In 1965 Lotfi Zadeh, Professor of University of Berkely, proposed a different notion of a set, in which elements can belong to a set to some extent and not definitively, as it is in case of the classical set theory. This proposal turned out applicable in many domains and initiated extensive research in fuzzy set theory, what became the name of Zadeh's theory [43].

In his approach an element can belong to a set to a degree k $(0 \leq k \leq 1)$, in contrast to classical set theory, where an element must definitely belong or not to a set. For example, in classical set theory one can say that someone is definitely ill or healthy, whereas in the fuzzy set theory language we can say that someone is ill (or healthy) at the 60 percent level (i.e., in degree 0.6).

Let us observe that the definition of fuzzy set involves more advanced mathematical concepts – real numbers and functions – whereas in classical set theory the notion of a set is used as a fundamental notion of whole mathematics and is used to derive any other mathematical concepts, e.g., numbers and functions. Consequently fuzzy set theory cannot replace classical set theory, because, in fact, the theory is needed to define fuzzy sets.

Fuzzy membership function has the following properties:

a) $\mu_{U-X}(x) = 1 - \mu_X(x)$ for any $x \in U$;
b) $\mu_{X \cup Y}(x) = max(\mu_X(x),\ \mu_Y(x))$ for any $x \in U$;
c) $\mu_{X \cap Y}(x) = min(\mu_X(x),\ \mu_Y(x))$ for any $x \in U$.

That means that the membership of an element to the union and intersection of sets is uniquely determined by its membership to constituent sets. This is a very nice property and allows very simple operations on fuzzy sets, which is a very important feature both theoretically and practically.

Several generalizations of this basic approach to concept approximation are presented in the literature (see, e.g., [14,42,44,45,50]).

Let us stress once more that classical set is a primitive notion and is defined intuitively or axiomatically. Fuzzy sets are defined by employing the fuzzy membership function, which involves advanced mathematical structures, numbers and functions. Thus it cannot play an analogous role in mathematics similar to that played by the classical concept of a set, which is used to define numbers and functions.

Fuzzy set theory can be perceived as new model of vagueness. The theory and its applications developed very extensively over the past four decades and attracted attention of engineers, logicians, mathematicians and philosophers worldwide.

6 Rough Sets

Data! data! data!
Sir Artur Conan Doyle (1994). The Adventures of Sherlock Holmes. Penguin Books, London

Rough set theory, proposed by the author in 1982 [28,30], is still another approach to vagueness.

Rough set theory expresses vagueness not by means of membership, but by employing a boundary region of a set. If the boundary region of a set is empty it means that the set is crisp, otherwise the set is rough (inexact). A nonempty boundary region of a set means that our knowledge about the set is not sufficient to define the set precisely.

In a manner similar to fuzzy set theory, rough set theory it is not an alternative to classical set theory but it is embedded in it. Rough set theory can be viewed as a specific implementation of Frege's idea of vagueness, i.e., imprecision in this approach is expressed by a boundary region of a set, and not by a partial membership, as in fuzzy set theory.

The rough set concept can be defined quite generally by means of topological operations, *interior* and *closure*, called *approximations.* At the onset of an introduction to rough sets, it was observed that the key to the presented approach is provided by the exact mathematical formulation of the concept of approximative (rough) equality of sets in a given approximation space [28]. In [30], an approximation space is represented by the pair (U, R), where U is a universe of objects, and $R \subseteq U \times U$ is an indiscernibility relation defined by an attribute set. The relation R is an equivalence relation. Let $[x]_R$ denote an equivalence class of an element $x \in U$ under the indiscernibility relation R, where $[x]_R = \{y \in U \mid xRy\}$.

In this context, R-approximations of any set $X \subseteq U$ are based on the exact (crisp) containment of sets. Then set approximations are defined as follows:

- $x \in U$ belongs with certainty to the R-lower approximation of $X \subseteq U$, if $[x]_R \subseteq X$.
- $x \in U$ belongs with certainty to the complement set of $X \subseteq U$, if $[x]_R \subseteq U - X$.
- $x \in U$ belongs with certainty to the R-boundary region of $X \subseteq U$, if $[x]_R \cap X \neq \oslash$ and $[x]_R \cap (U - X) \neq \oslash$.

Generalized approximation spaces were introduced in [42]. A *generalized approximation space* is a system $GAS = (U, I, \nu)$ where

- U is a non-empty set of objects, and $\mathcal{P}(U)$ is the powerset of U;
- $I : U \to \mathcal{P}(U)$ is an uncertainty function such that $x \in I(x)$ for any $x \in U$;
- $\nu : \mathcal{P}(\mathrm{U})$ x $\mathcal{P}(\mathrm{U}) \to [0, 1]$ denotes rough inclusion

The uncertainty function I defines for every object x a set of similarly defined objects. In effect, I defines a neighborhood of every sample element x belonging to the universe U (see, e.g., [32]). The rough inclusion function ν computes the degree of overlap between two subsets of U. Let $\mathcal{P}(U)$ denote the powerset of U. In general, rough inclusion $\nu : \mathcal{P}(\mathrm{U})$ x $\mathcal{P}(\mathrm{U}) \to [0, 1]$ can be defined in terms of the relationship between two sets, e.g., by

$$\nu(X, Y) = \begin{cases} \frac{|X \cap Y|}{|Y|}, & \text{if Y } \neq \emptyset \\ 1 \quad , & \text{otherwise} \end{cases}$$

for any $X, Y \subseteq U$.

From practical point of view it is better to define basic concepts of this theory in terms of data. Therefore we will start our considerations from a data set called an *information system*. An information system is a data table containing rows labeled by objects of interest, columns labeled by attributes and entries of the table are attribute values. For example, a data table can describe a set of patients in a hospital. The patients can be characterized by some attributes, like *age, sex, blood pressure, body temperature*, etc. With every attribute a set of its values is associated, e.g., values of the attribute age can be *young, middle*, and *old.* Attribute values can be also numerical. In data analysis the basic problem we are interested in is to find patterns in data, i.e., to find relationship between some set of attributes, e.g., we might be interested whether *blood pressure* depends on *age and sex.*

Let us describe this problem more precisely. Suppose we are given a finite, non-empty set of objects U called the *universe* and a set of *attributes* A, describing objects of the universe in terms of *attribute values.* Let X be a subset of U and B a subset of A. We want to characterize the set X in terms of attributes B. To this end we will need the basic concepts of rough set theory given below.

- The *lower approximation* of a set X with respect to B is the set of all objects, which can be for *certain* classified as X using B (are *certainly* X in view of B).
- The *upper approximation* of a set X with respect to B is the set of all objects which can be *possibly* classified as X using B (are *possibly* X in view of B).
- The *boundary region* of a set X with respect to B is the set of all objects, which can be classified neither as X nor as not-X using B.

Now we are ready to give the definition of rough sets.

- Set X is *crisp* (exact with respect to B), if the boundary region of X is empty.
- Set X is *rough* (inexact with respect to B), if the boundary region of X is nonempty.

Thus a set is *rough* (imprecise) if it has nonempty boundary region; otherwise the set is crisp (precise). This is exactly the idea of vagueness proposed by Frege.

Let us observe that the definition of rough sets refers to data (knowledge), and is *subjective*, in contrast to the definition of classical sets, which is in some sense an *objective* one.

The approximations and the boundary region can be defined more precisely. To this end we need some additional notation.

Every subset of attributes B determines an equivalence relation on U. This relation will be referred to as an *indiscernibility relation.* The equivalence class determined by an element x and the set of attributes B will be denoted $B(x)$. The indiscernibility relation in a certain sense describes our lack of knowledge about the universe. Equivalence classes of the indiscernibility relation, called *granules* generated by the set of attributes B, represent an elementary portion of knowledge we are able to perceive in terms of available data. Thus in view of the data we are unable, in general, to observe individual objects but we are forced to reason only about the accessible granules of knowledge (see, e.g., [27,30,35].

Formal definitions of approximations and the boundary region are as follows:

- *B-lower approximation* of X

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

- *B-upper approximation* of X

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

- *B-boundary region* of X

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

As we can see from the definition approximations are expressed in terms of granules of knowledge. The lower approximation of a set is union of all granules determined by the set of attributes B which are entirely included in the set; the upper approximation is union of all granules which have non-empty intersection with the set; the boundary region of set is the difference between the upper and the lower approximation.

Thus the definition of rough set also requires advanced mathematical concepts (relations) and consequently, similarly as fuzzy set, cannot replace classical concept of a set.

Several generalizations of the above approach have been proposed in the literature (see, e.g., [14,27,42,44,45,50]). In particular, in some of these approaches the set inclusion to a degree is used instead of the exact inclusion. It is worthwhile to mention that the set inclusion to a degree has been considered by Łukasiewicz [20] in studies on assigning fractional truth values to logical formulas.

Different aspects of vagueness in the roughs set framework are discussed, e.g., in [21,24,36,41].

Our knowledge about the approximated concepts is often partial and uncertain [15]. For example, the concept approximation should be constructed from examples and counter examples of objects for the concepts [16]. Hence, the concept approximations constructed from a given sample of objects is extended, using inductive reasoning, on unseen so far objects. The rough set approach for dealing with concept approximation under such partial knowledge is presented, e.g., in [44]. Moreover, the concept approximations should be constructed under dynamically changing environments [41]. This leads to a more complex situation when the boundary regions are not crisp sets what is consistent with the postulate of the higher order vagueness, considered by philosophers (see, e.g., [17]). It is worthwhile to mention that is has been also developed a rough set approach to approximation of compound concepts that we are unable to approximate using the traditional methods [7,47]. The approach is based on hierarchical learning and ontology approximation [27,22,43,5] (see Section 8). Approximation of concepts in distributed environments is discussed in [40]. A survey of algorithmic methods for concept approximation based on rough sets and boolean reasoning in presented in [39].

7 Fuzzy Versus Rough

In order to compare both concepts, fuzzy and rough sets we let us observe that rough sets can be also defined employing, instead of approximation, *rough membership function* [31]

$$\mu_X^B : U \to < 0, 1 >,$$

where

$$\mu_X(x) = \frac{card(B(x) \cap X)}{card(X)},$$

and $card(X)$ denotes the cardinality of X.

The rough membership function expresses a conditional probability that x belongs to X given B and can be interpreted as a degree that x belongs to X in view of knowledge about x expressed by B. This means that the definition reflects a subjective knowledge about elements of the universe, in contrast to classical definition of a set.

It can be shown that the rough membership function has the following properties [31]:

1) $\mu_X^B(x) = 1$ *iff* $x \in B_*(X)$;
2) $\mu_X^B(x) = 0$ *iff* $x \in U - B^*(X)$;
3) $0 < \mu_X^B(x) < 1$ *iff* $x \in BN_B(X)$;
4) $\mu_{U-X}^B(x) = 1 - \mu_X^B(x)$ *for any* $x \in U$;
5) $\mu_{X \cup Y}^B(x) \geq max(\mu_X^B(x),\ \mu_Y^B(x))$ *for any* $x \in U$;
6) $\mu_{X \cap Y}^B(x) \leq min(\mu_X^B(x),\ \mu_Y^B(x))$ *for any* $x \in U$.

From the properties it follows that the rough membership differs essentially from the fuzzy membership, for properties 5) and 6) show that the membership for union and intersection of sets, in general, cannot be computed – as in the case of fuzzy sets – from their constituents membership. Thus formally the rough membership is more general from fuzzy membership. Moreover, the rough membership function depends on an available knowledge (represented by attributes from B). Besides, the rough membership function, in contrast to fuzzy membership function, has a probabilistic flavor.

Let us also mention that rough set theory, in contrast to fuzzy set theory, clearly distinguishes two very important concepts, vagueness and uncertainty, very often confused in the AI literature. Vagueness is the property of sets and can be described by approximations, whereas uncertainty is the property of elements of a set and can be expressed by the rough membership function.

Both fuzzy and rough set theory represent two different approaches to vagueness. Fuzzy set theory addresses *gradualness* of knowledge, expressed by the fuzzy membership – whereas rough set theory addresses *granularity* of knowledge, expressed by the indiscernibility relation. A nice illustration of this difference has been given by Dider Dubois and Henri Prade [11] in the following example. In image processing fuzzy set theory refers to gradualness of gray level, whereas rough set theory is about the size of pixels.

Consequently, both theories are not competing but are rather complementary. In particular, the rough set approach provides tools for approximate construction of fuzzy membership functions. The rough-fuzzy hybridization approach proved to be successful in many applications (see, e.g., [25,26]).

Interesting discussion of fuzzy and rough set theory in the approach to vagueness can be found in [36].

Finally, let us observe that fuzzy set and rough set theory are not a remedy for classical set theory difficulties.

8 Logic and Rough Sets

Reality, or the world we all know, is only a description.
Carlos Castaneda (1972). Journey to Ixtlan: The lesson of Don Juan.
Simon & Schuster, New York

The father of contemporary logic is a German mathematician Gottlob Frege (1848-1925). He thought that mathematics should not be based on the notion of set but on the notions of logic. He created the first axiomatized logical system but it was not understood by the logicians of those days.

In the thirties of the previous century a rapid development of logic took place, to which Polish logicians contributed to a large extent, in particular Alfred Tarski (1901-1983).

Development of computers and their applications stimulated logical research and widened their scope.

When we speak about logic we generally mean *deductive logic*. It gives us tools designed for deriving true propositions from other true propositions. Deductive reasoning always leads to true conclusions. The theory of deduction has well established generally accepted theoretical foundations. Deductive reasoning is the main tool used in mathematical reasoning and found no application beyond it.

Rough set theory has contributed to some extent to various kind of deductive reasoning. Particularly, rough set methodology contributed essentially to modal logics, many valued logic, intuitionistic logic and others (see, e.g., [3]). A summary of this research can be found in [33] and interested reader is advised to consult this volume.

In natural sciences (e.g., in physics) *inductive reasoning* is of primary importance. The characteristic feature of such reasoning is that it does not begin from axioms (expressing general knowledge about the reality) like in deductive logic, but some partial knowledge (examples) about the universe of interest are the starting point of this type of reasoning, which are generalized next and they constitute the knowledge about wider reality than the initial one. In contrast to deductive reasoning, inductive reasoning does not lead to true conclusions but only to probable (possible) ones. Also in contrast to the logic of deduction, the logic of induction does not have uniform, generally accepted, theoretical foundations as yet, although many important and interesting results have been obtained, e.g., concerning statistical and computational learning and others.

Verification of validity of hypotheses in the logic of induction is based on experiment rather than the formal reasoning of the logic of deduction. Physics is the best illustration of this fact.

The research on inductive logic have a few centuries', long history and outstanding English philosopher John Stuart Mill (1806-1873) is considered its father.

The creation of computers and their innovative applications essentially contributed to the rapid growth of interest in inductive reasoning. This domain develops very dynamically thanks to computer science. Machine learning, knowledge discovery, reasoning from data, expert systems and others are examples of new directions in inductive reasoning. It seems that rough set theory is very well suited as a theoretical basis for inductive reasoning. Basic concepts of this theory fit very well to represent and analyze knowledge acquired from examples, which can be next used as starting point for generalization. Besides, in fact rough set theory has been successfully applied in many domains to find patterns in data (data mining) and acquire knowledge from examples (learning from examples). Thus, rough set theory seems to be another candidate as a mathematical foundation of inductive reasoning [5,22,44].

The most interesting from computer science point of view is *common sense* reasoning. We use this kind of reasoning in our everyday life, and examples of such kind of reasoning we face in news papers, radio TV etc., in political, economic etc., debates and discussions.

The starting point to such reasoning is the knowledge possessed by the specific group of people (*common knowledge*) concerning some subject and intuitive methods of deriving conclusions from it. We do not have here possibilities of resolving the dispute by means of methods given by deductive logic (reasoning) or by inductive logic (experiment). So the best known methods of solving the dilemma is voting, negotiations or even war. See e.g., Gulliver's Travels [46], where the hatred between Tramecksan (High-Heels) and Slamecksan (Low-Heels) or disputes between Big-Endians and Small-Endians could not be resolved without a war.

These methods do not reveal the truth or falsity of the thesis under consideration at all. Of course, such methods are not acceptable in mathematics or physics. Nobody is going to solve by voting, negotiations or declare a war – the truth of Fermat's theorem or Newton's laws.

Reasoning of this kind is the least studied from the theoretical point of view and its structure is not sufficiently understood, in spite of many interesting theoretical research in this domain [13]. The meaning of common sense reasoning, considering its scope and significance for some domains, is fundamental and rough set theory can also play an important role in it but more fundamental research must be done to this end [43].

In particular, the rough truth introduced in [29] and studied, e.g., in [2,4] seems to be important for investigating commonsense reasoning in the rough set framework.

Let us consider a simple example. In the considered decision table we assume $U = Birds$ is a set of birds that are described by some condition attributes from a set A. The decision attribute is a binary attribute $Flies$ with possible values yes if the given bird flies and no, otherwise. Then, we define the set of abnormal birds by $Ab_A(Birds) = A_*(\{x \in Birds : Flies(x) = no\})$. Hence, we have, $Ab_A(Birds) = Birds - A^*(\{x \in Birds : Flies(x) = yes\})$ and $Birds - Ab_A(Birds) = A^*(\{x \in Birds : Flies(x) = yes\})$. It means that for normal birds it is consistent, with knowledge represented by A, to assume that they can fly, i.e., it is possible that they can fly. One can optimize $Ab_A(Birds)$ using A to obtain minimal boundary region in the approximation of $\{x \in Birds : Flies(x) = no\}$.

It is worthwhile to mention that in [10] has been presented an approach combining the rough sets with nonmonotonic reasoning. There are distinguished some basic concepts that can be approximated on the basis of sensor measurements and more complex concepts that are approximated using so called transducers defined by first order theories constructed overs approximated concepts. Another approach to commonsense reasoning has been developed in a number of papers (see, e.g., [35,43,22,27,5]). The approach is based on an ontological framework for approximation. In this approach approximations are constructed for concepts and dependencies between the concepts represented in a given ontology expressed, e.g., in natural language. Still another approach combining rough sets with logic programming is discussed in [48].

To recapitulate, the characteristics of the three above mentioned kinds of reasoning are given below:

1. deductive:
   - reasoning method: axioms and rules of inference;
   - applications: mathematics;
   - theoretical foundations: complete theory;
   - conclusions: true conclusions from true premisses;
   - hypotheses verification: formal proof.
2. inductive:
   - reasoning method: generalization from examples;
   - applications: natural sciences (physics);
   - theoretical foundation: lack of generally accepted theory;
   - conclusions: not true but probable (possible);
   - hypotheses verification - experiment.
3. common sense:
   - reasoning method based on common sense knowledge with intuitive rules of inference expressed in natural language;
   - applications: every day life, humanities;

– theoretical foundation: lack of generally accepted theory;
– conclusions obtained by mixture of deductive and inductive reasoning based on concepts expressed in natural language, e.g., with application of different inductive strategies for conflict resolution (such as voting, negotiations, cooperation, war) based on human behavioral patterns;
– hypotheses verification - human behavior.

9 Conclusions

Basic concept of mathematics, the set, leads to antinomies, i.e., it is contradictory.

The deficiency of sets, has philosophical rather than practical meaning, since sets used in mathematics are free from the above discussed faults. Antinomies are associated with very "artificial" sets constructed in logic but not found in sets used in mathematics. That is why we can use mathematics safely.

Fuzzy set and rough set theory are two different approaches to vagueness and are not remedy for classical set theory difficulties.

Fuzzy set theory addresses gradualness of knowledge, expressed by the fuzzy membership - whereas rough set theory addresses granularity of knowledge, expressed by the indiscernibility relation.

From practical point of view both theories are not competing but are rather complementary.

Summing up:

– The notion of classical set is fundamental for whole mathematics and is necessary to provide rigor in mathematics.
– Non-classical sets (fuzzy and rough) cannot replace classical sets - for their definitions need classical set theory (i.e., more advanced mathematical concepts, real numbers, functions and relations).
– The classical sets, lead to antinomies.
– The deficiency of classical sets has rather philosophical than practical meaning, since sets used in everyday mathematics are free from antinomies.
– Non-classical sets (fuzzy and rough) are not remedy for classical set theory difficulties but are two different approaches to vagueness.

Summary

In this paper a brief discussion on the rough set concept and its place in various ideas of sets is presented. The article is not intended to serve as an introduction to rough set theory but is rather meant to give some philosophical background underlining the theory.

Acknowledgments

Thanks are due to Professor Andrzej Skowron and Professor Jim Peters for critical remarks.

References

1. P. Apostoli, A. Kanda (1999). *Parts of the Continuum: Towards a Modern Ontology of Sciences.* Technical Reports in Philosophical Logic **96**(1), **97**(1), Revised March, 1999, The University of Toronto, Department of Philosophy.
2. M. Banerjee, M.K. Chakraborty (1994). Rough Consequence and Rough Algebra. In: W.P. Ziarko (ed.), *Rough Sets, Fuzzy Sets and Knowledge Discovery, Proc. Int. Workshop on Rough Sets and Knowledge Discovery (RSKD'93)*, Workshops in Computing, Springer-Verlag & British Computer Society, London, Berlin, 196–207.
3. M. Banerjee, M.K. Chakraborty (2004). Algebras from Rough Sets. In: [27], 157–188.
4. M. Banerjee (2004). Rough Truth, Consequence, Consistency and Belief Revision. In: S. Tsumoto et. al. (eds.), *Proc. 4th Int. Conf. On Rough Sets and Current Trends in Computing (RSCTC2004)*, Uppsala, Sweden, June 2004, LNAI **3066**, Springer-Verlag, Heidelberg, 95–102.
5. J.G. Bazan, J.F. Peters, A. Skowron (2005). Behavioral Pattern Identification Through Rough Set Modelling. *Proceedings of RSFDGrC 2005*, Regina, Canada, September 1–3, 2005, LNCS, Springer, Heidelberg (to appear).
6. M.K. Chakraborty, M. Banerjee (1993). Rough Consequence. *Bull. Polish Acad. Sc. (Math.)*, **41**(4): 299–304.
7. L. Breiman (2001). Statistical Modeling: The Two Cultures. *Statistical Science* **16**(3): 199–231.
8. G. Cantor (1883). *Grundlagen einer allgemeinen Mannigfaltigkeitslehre.* Leipzig, 1883.
9. R. Casati, A. Varzi (1999). *Parts and Places. The Structures of Spatial Representation, Cambridge, MA, and London: MIT Press, Bradford Books.*
10. P. Doherty, W. Łukaszewicz, A. Skowron, A. Szałas (2005). *Knowledge Engineering: A Rough Set Approach.* Springer, Heidelberg (to appear).
11. D. Dubois, H. Prade (1991). Foreword. In: Zdzisław Pawlak, *Rough Sets, Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht.
12. G. Frege (1893). *Grundlagen der Arithmetik* **2**, Verlag von Herman Pohle, Jena.
13. D.M. Gabbay, C.J. Hogger, J.A. Robinson (eds.) (1994). *Handbook of Logic in Aretificial Intelligence and Logic Programming: Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*, Calderon Press, Oxford.
14. S. Greco, B. Matarazzo, R. Słowiński, (2001). Rough Set Theory for Multicriteria Decision Analysis. *European Journal of Operational Research* **129**(1): 1–47.
15. J.W. Grzymała-Busse (1990). *Managing Uncertainty in Expert Systems.* Kluwer Academic Publishers, Norwell, MA.
16. J. Friedman, T. Hastie, R. Tibshirani (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag, Heidelberg.
17. R. Keefe (2000). *Theories of Vagueness.* Cambridge Studies in Philosophy, Cambridge, UK.
18. R. Keefe, P. Smith (eds.) (1997). *Vagueness: A Reader.* MIT Press, Massachusetts, MA.
19. S. Leśniewski (1929). Grungzüge eines neuen Systems der Grundlagen der Mathematik. *Fundamenta Matemaicae* **14**: 1–81.
20. J. Łukasiewicz (1913). Die Logischen grundlagen der Wahrscheinlichkeitsrechnung. Kraków, 1913, In: L. Borkowski (ed.), *Jan Łukasiewicz - Selected Works*, North Holland Publishing Company, Amsterdam, London, Polish Scientific Publishers, Warsaw, 1970.

21. S. Marcus (1998). The Paradox of the Heap of Grains, in Respect to Roughness, Fuzziness and Negligibility. In L.Polkowski, A. Skowron (eds.) *First International Conference on Rough Sets and Current Trends in Computing* (RSCTC'98), Warsaw, Poland, June 22-26, 1998, LNAI **1424**, Springer-Verlag, Heidelberg, 19–23.
22. S. Hoa Nguyen, J. Bazan, A. Skowron, and H. Son Nguyen (2004). Layered Learning for Concept Synthesis. *Transactions on Rough Sets I: Journal Subline*, LNCS **3100**, Springer, Heidelberg, 187–208.
23. E. Orłowska (1984). Semantics of Vague Conepts. In: G. Dorn, P. Weingartner (eds.), *Foundation of Logic and Linguistics*, Plenum Press, New York, 465–482.
24. E. Orłowska (1987). *Reasoning about Vague Concepts. Bull. Polish Acad. Sci., Math.* **35**, 643–652.
25. S.K. Pal, A. Skowron (1999). *Rough Fuzzy Hybridization: A New Trend in Decision-Making.* Springer-Verlag, Singapore.
26. S.K. Pal, P. Mitra (2004). *Pattern Recognition Algorithms for Data Mining.* CRC Press, Boca Raton, Florida.
27. S.K. Pal, L. Polkowski, A. Skowron (Eds.) (2004). *Rough-Neural Computing: Techniques for Computing with Words.* Springer, Heidelberg, 2004.
28. Z. Pawlak (1982). Rough Sets. *Int. J. of Information and Computer Sciences* **11**(5), 341–356.
29. Z. Pawlak (1987). Rough Logic. *Bull. Polish. Acad. Sci., Tech.* **35**(5-6), 253–258.
30. Z. Pawlak (1991). *Rough Sets: Theoretical Aspects of Reasoning about Data.* System Theory, Knowledge Engineering and Problem Solving **9**, Kluwer Academic Publishers, Dordrecht.
31. Z. Pawlak, A. Skowron (1994). Rough Membership Functions. In: R. R. Yager, M. Fedrizzi, J. Kacprzyk (eds.), *Advances in the Dempster-Schafer Theory of Evidence*, John Wiley and Sons, New York, 251–271.
32. J.F. Peters, A, Skowron, P. Synak, S. Ramanna (2003). Rough Sets and Information Granulation. In: Bilgic, T., Baets, D., Kaynak, O. (eds.), *Tenth International Fuzzy Systems Association World Congress IFSA*, Istanbul, Turkey, June 30-July 2, 2003, LNAI **2715**, Springer-Verlag, Heidelberg, 370–377.
33. L. Polkowski (2002). *Rough Sets: Mathematical Foundations.* Physica-Verlag, Heidelberg.
34. L. Polkowski, A. Skowron (1996). Rough Mereology: A New Paradigm for Approximate Reasoning. *International Journal of Approximate Reasoning* **15**, 333–365.
35. L. Polkowski, A. Skowron (2001). Rough Mereological Calculi of Granules: A Rough Set Approach to Computation. *Computational Intelligence* **17**, 472–492.
36. S. Read (1995). *Thinking about Logic. An Introduction to the Philosophy of Logic.* Oxford University Press, Oxford, New York.
37. B. Russell (1903). *The Principles of Mathematics.* London, George Allen & Unwin Ltd., 1st Ed. 1903 (2nd Edition in 1937).
38. B. Russell (1940). *An Inquiry into Meaning and Truth.* George Allen and Unwin, London; W.W. Norton, New York.
39. A. Skowron (2000). Rough Sets in KDD. (plenary lecture) In: Z. Shi, B. Faltings, and M. Musen (eds.), *16-th World Computer Congress (IFIP'2000): Proceedings of Conference on Intelligent Information Processing (IIP'2000)*, Publishing House of Electronic Industry, Beijing, 1–17.
40. A. Skowron (2004). Approximate Reasoning in Distributed Environments. In: Zhong, N., Liu, J. (Eds.), *Intelligent Technologies for Information Analysis*, Springer, Heidelberg, 433-474.
41. A. Skowron (2005). Rough Sets and Vague Concepts. *Fundamenta Informaticae* **64**(1-4): 417–431.

42. A. Skowron, J. Stepaniuk (1996). Tolerance Approximation Spaces. *Fundamenta Informaticae* **27**(2-3): 245–253.
43. A. Skowron, J. Peters (2003). Rough Sets: Trends and Challenges (plenary talk). In: G. Wang, Q. Liu, Y.Y. Yao, A. Skowron (eds.), *Proceedings of the Ninth International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC'2003)*, Chongqing, China, May 26-29, 2003, LNAI **2639**, 25–34.
44. A. Skowron, R. Swiniarski, P. Synak (2005). Approximation Spaces and Information Granulation. *Transactions on Rough Sets III: Journal Subline*, LNCS **3400**, Springer, Heidelberg 2005, 175–189.
45. R. Słowiński, D. Vanderpooten (1997). Similarity Relation as a Basis for Rough Approximations. In: P. Wang (ed.), *Advances in Machine Intelligence and Soft Computing* **4**, Duke University Press, 17-33.
46. J. Swift (1726). *Gulliver's Travels into Several Remote Nations of the World*. London, M, DCC, XXVI.
47. V. Vapnik (1998). *Statistical Learning Theory.* John Wiley & Sons, New York, NY.
48. A. Vitória (2005). A Framework for Reasoning with Rough Sets. Licentiate Thesis, Linköping University 2004. *Transactions on Rough Sets IV: Journal Subline*, LNCS, Springer, Heidelberg (to appear).
49. P. Vopenka (1979). *Mathematics in the Alternative Set Theory*. Teubner, Leipzig.
50. W. Ziarko (1993). Variable Precision Rough Set Model. *Journal of Computer and System Sciences* **46**, 39–59.
51. L.A. Zadeh (1965). Fuzzy Sets. *Information and Control* **8**, 338–353.