AI AND INTELLIGENT INDUSTRIAL APPLICATIONS
THE ROUGH SET PERSPECTIVE

Zdzisław Pawlak

Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Poland, e-mail: zpw@ii.pw.edu.pl
Mailing address to correspondence: Institute of Computer Science, Warsaw University of Technology, 00665 Warsaw, ul. Nowowiejska 15/19, Poland, phon: +48 22 8345659, fax: +48 22 8251635

ABSTRACT
Application of intelligent methods in industry become a very challenging issue nowadays and will be of extreme importance in the future. Intelligent methods include, fuzzy sets neural networks genetics algorithms and others techniques known as soft computing. No doubt rough set theory can also contribute essentially to this domain. In this paper basic ideas of rough set theory are presented and some possible intelligent industrial applications outlined.

INTRODUCTION
Rough set theory is a new mathematical approach to data analysis. Basic idea of this method hinges on classification of objects of interest into similarity classes (clusters) containing objects which are indiscernible with respects to some features, e.g., colour, temperature etc., which form basic building blocks of knowledge about reality, and are employed next to find out hidden patterns in data. Basis of rough set theory can be found in Pal & Skowron (1999), Pawlak (1991), Pawlak et al., (1995) and Polkowski & Skowron (1998).

Rough set theory has some overlaps with other methods of data analysis, e.g., statistics, cluster analysis, fuzzy sets, evidence theory and other but it can be viewed in its own rights as an independent discipline.

The rough set approach seems to be of fundamental importance to AI and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, inductive reasoning and pattern recognition. It seems of particular importance to decision support systems and data mining.

Rough set theory has been successfully applied in many real-life problems in medicine, pharmacology, engineering, banking, financial and market analysis and others. More about applications of rough set theory can be found in Lin & Cicerone (1997), Polkowski & Skowron (1998), Slowinski (1992), Sładowski & Ziarko (1993), Tsumoto et al. (1996), Wang (1997) and Ziarko (1993) and others.

Very promising new areas of application of the rough set concept seems to emerge in the near future. They include rough control, rough data bases, rough information retrieval, rough neural network and others.

ROUGH SETS AND INTELLIGENT INDUSTRIAL APPLICATIONS
Artificial intelligence approach to industrial process is real challenge for industry in the years to come. Rough set theory seems to be particularly suited for problem solving in this area. Some of them are briefly discussed below.

1) Material sciences. Application of rough sets to new materials design and investigating material properties has already shown its usefulness in this area. Pioneer work in this domain
is due to Jackson et al. (1994, 1996). It is interesting also to mention in this context works on application of rough sets to investigation of the relationship between structure and activity of drugs (Krysiński, 1995). The method used here can be also used not only in the case of drugs but for any other kind of materials.


3) **Decision support systems.** Rough set based decision support systems can be widely used in many kinds of industrial decision making on various levels, stretching down from specific industrial process up to management and business decisions (Golan & Ziarko, 1995, Pawlak, 1994, Słowiński, 1992, 1995, Stepniak, 1996 and Ziarko et al., 1993).

4) **Machine diagnosis.** Rough set approach has been used to technical diagnosis of mechanical objects (Nowicki et al., 1990, 1992, Słowiński et al., 1996, Słowiński & Zopounidis, 1995 and Stefanowski et al., 1992).

5) **Neural networks.** Neural networks have found many interesting applications in intelligent control of industrial processes. Combining neural networks with fuzzy sets adds new dimension to this domain. Rough sets and neural networks can be also linked together and give better results and greater speed then the classical neural network approach alone (Lingras, 1996, Mitra & Banerjee, 1996, Nguyen et al., 1997, Nguyen et al., 1995 and Szczuka, 1996).

6) **Varia.** Beside the above said domains of intelligent industrial applications of rough sets there are many other fields where rough set approach can be useful. They include expert systems (Chen et al., 1997), engineering design (Arciszewski & Ziarko, 1987, 1990), signal and image processing (Kowalczyk, 1996), data bases and information retrieval (Baubouef et al., 1995, Funakoshi & Tu Bao Ho, 1996) and others (An et al., 1997, Furuta et al., 1996, Rubin et al., 1996 and Zak & Stefanowski, 1994).

The above discussed list of possible application of rough sets is of course not exhaustive one but shows areas where application of rough set have already proved to be of use.

Rough sets approach shows many advantages. The most important ones are listed below.

- Provides efficient algorithms for finding hidden patterns in data.
- Identifies relationships that would not be found using statistical methods.
- Allows both qualitative and quantitative data.
- Finds minimal sets of data (data reduction).
- Evaluates significance of data.
- Generates sets of decision rules from data.
- It is easy to understand.
- Offers straightforward interpretation of obtained results.
No doubt rough set theory can be very useful in many branches of intelligent industrial applications as a independent, complementary approach or combined together with other areas of soft computing, e.g. fuzzy sets, neural networks, etc.

**APPROXIMATIONS – BASIC CONCEPTS OF ROUGH SET THEORY**

Data are usually given in a form of a data table, called also an attribute-value table, an information table or a database. Data table is a matrix rows of which are labelled by objects, whereas columns are labelled by attributes. Entries of the table are attribute values. An example of a database is shown in Table 1.

<table>
<thead>
<tr>
<th>Store</th>
<th>E</th>
<th>Q</th>
<th>L</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>high</td>
<td>good</td>
<td>no</td>
<td>profit</td>
</tr>
<tr>
<td>2</td>
<td>med.</td>
<td>good</td>
<td>no</td>
<td>loss</td>
</tr>
<tr>
<td>3</td>
<td>med.</td>
<td>good</td>
<td>no</td>
<td>profit</td>
</tr>
<tr>
<td>4</td>
<td>no</td>
<td>avg.</td>
<td>no</td>
<td>loss</td>
</tr>
<tr>
<td>5</td>
<td>med.</td>
<td>avg.</td>
<td>yes</td>
<td>loss</td>
</tr>
<tr>
<td>6</td>
<td>high</td>
<td>avg.</td>
<td>yes</td>
<td>profit</td>
</tr>
</tbody>
</table>

Table 1. Example of a database

In the database six stores are characterized by four attributes:

- $E$ – empowerment of sales personnel,
- $Q$ – perceived quality of merchandise,
- $L$ – high traffic location,
- $P$ – store profit or loss.

Suppose we are interested which features are associated with profit or loss of stores. This problem cannot be solved uniquely since the data are inconsistent, i.e., stores 2 and 3 have the same features (values of attributes $E$, $Q$, $L$) but store 2 has loss, whereas store 3 has profit. The situation is depicted in Fig. 1.

![Figure 1. Classification of stores](image-url)
Usually to avoid this kind of inconsistency a probabilistic or fuzzy characterization of stores is assumed. Also adding more attributes can resolve the inconsistency. In the rough set theory the approach is different. We try to preserve the data intact and find other ways out, and see what the original data is telling us. To this end we propose instead probabilistic or fuzzy methods to use topological methods of reasoning about the data, by introducing the concepts of the lower and the upper approximation of sets, which in fact are topological interior and closure operations. Thus we can distinguish the following classes of stores:

- the set \{1, 3, 6\} of all stores having \textit{profit},
- the set \{1, 6\} of all stores \textit{certainly} having \textit{profit} (the \textit{lower approximation} of the set \{1, 3, 6\}),
- the set \{1, 2, 3, 6\} of all stores \textit{possibly} having \textit{profit} (the upper approximation of the set \{1, 3, 6\}),
- the set \{2, 3\} of all stores that can be classified as having neither \textit{profit} nor \textit{loss} (the \textit{boundary region} of the set \{1, 3, 6\}),
- the set \{2, 4, 5\} of all stores having \textit{loss},
- the set \{4, 5\} of all stores \textit{certainly} having \textit{loss} (the \textit{lower approximation} of the set \{2, 4, 5\}),
- the set \{2, 3, 4, 5\} of all stores possibly having \textit{loss} (the \textit{upper approximation} of the set \{2, 4, 5\}),
- the set \{2, 3\} of all stores that can be classified as having neither \textit{profit} nor \textit{loss} (the \textit{boundary region} of the set \{2, 4, 5\}).

Let us examine now the case more closely.

Each subset of attributes in the data table determines a partition all objects into clusters having the same attribute values, or in other word displaying the same features expressed in terms of attribute values. In other words all objects revealing the same features are \textit{indiscernible} (similar) in view of the available information and form blocks, which can be understood as elementary granules of knowledge. These granules are called \textit{elementary sets} or \textit{concepts}, and can be considered as elementary building blocks (atoms) of our knowledge about reality we are interested in. Elementary concepts can be combined into \textit{compound concepts}, i.e. concepts that are uniquely determined in terms of elementary concepts. Any union of elementary sets is called a crisp set, and any other sets are referred to as rough (vague, imprecise). With every set \(X\) we can associate two crisp sets, called the \textit{lower} and the \textit{upper approximation} of \(X\). The lower approximation of \(X\) is the union of all elementary set which are included in \(X\), whereas the upper approximation of \(X\) is the union of all elementary set which have non-empty intersection with \(X\). In other words the lower approximation of a set is the set of all elements that \textit{surely} belongs to \(X\), whereas the upper approximation of \(X\) is the set of all elements that \textit{possibly} belong to \(X\). The difference of the upper and the lower approximation of \(X\) is its \textit{boundary region}. Obviously a set is rough if it has non empty boundary region whatsoever; otherwise the set is crisp. Elements of the boundary region can be classified, employing the available knowledge, neither to the set nor its complement. Approximations of sets are basic operations in rough set theory and are used as main tools to deal with vague and uncertain data.

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 a database. 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:

$$(x, y) \in I(B) \text{ if and only if } a(x) = a(y) \text{ for every } a \in A,$$

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 $U/B$; 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.

The indiscernibility relation will be used next to define two basic operations in rough set theory, which are defined below:

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

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

and are called the $B$-lower and the $B$-upper approximation of $X$, respectively. The set

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

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

For example:

assuming $B = \{E, Q, L\}$ the $B$-lower approximation of the set $X_{profit} = \{1, 3, 6\}$ in the set $B^*(X_{profit}) = \{1, 6\}$, the $B$-upper approximation – is the set $B^*(X_{profit}) = \{1, 2, 3, 6\}$, whereas the set $BN_B(X_{profit}) = \{2, 3\}$ is boundary region of the set $X_{profit} = \{1, 3, 6\}$.

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

Summing up:

- 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 $X$),
- 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 $X$),
- 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 $B$.

Figure 2 gives graphical illustrates of approximations and the boundary region.
Granules of knowledge
The set of objects

The lower approximation
The set
The upper approximation

Figure 2. Approximations and the boundary region

THE MEMBERSHIP FUNCTION

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

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

Obviously

$$0 \leq \mu_X^B(x) \leq 1.$$  

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

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 has the following properties:

a) $\mu_X^B(x) = 1$ iff $x \in B_+(X)$,
b) $\mu_X^B(x) = 0$ iff $x \in -B^*(X)$,
c) $0 < \mu_X^B(x) < 1$ iff $x \in BN^B_-(X)$,
d) If $I(B) = \{(x,x) : x \in U\}$, then $\mu_X^B(x)$ is the characteristic function of $I(B)$,
e) If $(x,y) \in I(B)$, then $\mu_X^B(x) = \mu_X^B(y)$ provided $I(B)$,
f) $\mu_{U^c-X}^B(x) = 1 - \mu_X^B(x)$ for any $x \in U$,
\( g \) \) \( \mu_{X \cup Y}(x) \geq \max(\mu_X^B(x), \mu_Y^B(x)) \) for any \( x \in U \),

\( h \) \( \mu_{X \cup Y}^B(x) \leq \min(\mu_X^B(x), \mu_Y^B(x)) \) for any \( x \in U \),

\( i \) If \( X \) is a family of pair wise disjoint sets of \( U \), then \( \mu_{X \cup Y}^B(x) = \sum_{x \in X} \mu_X^B(x) \) for any \( x \in U \).

The above properties show clearly the difference between fuzzy and rough memberships. In particular properties \( g \) and \( h \) show that the rough membership can be regarded formally as a generalization of fuzzy membership, for the max and the min operations for union and intersection of sets respectively for fuzzy sets are special cases of that for rough sets. But let us recall that the "rough membership", in contrast to the "fuzzy membership", has probabilistic flavour.

It can be easily seen that 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. Rough set approach shows clear connection between these two concepts.

**DEPENDENCY OF ATTRIBUTES**

Approximations of sets are strictly related with the concept of dependency (total or partial) of attributes.

Often we distinguish in a database two sets of attributes, called condition and decision attributes. For example, in Table 1 \( E, Q, L \) are condition and \( P \) is the decision attribute.

Intuitively, a set of decision attributes \( D \) depends totally on a set of condition attributes \( C \), denoted \( C \Rightarrow D \), if all values of attributes from \( D \) are uniquely determined by value of attribute form \( C \). In other words, \( D \) depends totally on \( C \), if there exists a functional dependency between values of \( D \) and \( C \).

We would also need a more general concept of dependency of attributes, called the partial dependency of attributes. 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

\[
D = \{ \gamma(C, D) \} = \sum_{x \in U/D} \frac{\text{card}(C(x))}{\text{card}(U)}
\]

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 block of the partition \( U/D \), employing attributes \( C \) and will be called the degree of the dependency, which can be also interpreted as a probability that \( x \in U \) belongs to one of the decision classed determined by decision attributes.

The degree of dependency between the set of attributes \( \{E, Q, L\} \) and the attribute \( P \) is \( 2/3 \).

**REDUCTION OF ATTRIBUTES**

We often face a question whether we can remove some data from a data-table preserving its basic properties, that is – whether a table contains some superfluous data.

Let us express this idea more precisely.
Let $C, D \subseteq A$, be sets of condition and decision attributes, respectively. We will say that $C' \subseteq C$ is a $D$-reduct (reduct with respect to $D$) of $C$, if $C'$ is a minimal subset of $C$ such that 

$$\gamma(C, D) = \gamma(C', D).$$

Hence any reduct enables us to reduce condition attributes in such a way that the degree of dependency between condition and decision attributes is preserved. In other words reduction of condition attributes removes superfluous conditions attributes and gives a minimal number of conditions necessary to make specified decisions.

For example, the set of attributes $\{E, Q, L\}$ has two reducts $\{E, Q\}$ and $\{E, L\}$.

**DECISION RULES**

Every dependency $C \Rightarrow_D D$ can be described by a set of decision rules in the form „if ... then”. Decision rules are implications $\Phi \rightarrow \Psi$, where $\Phi$ and $\Psi$ are formulas called conditions and decisions of the rule respectively – built up from elementary formulas (attribute, value) combined together by means of propositional connectives „and”, „or” and „not” in a standard way.

An example of a decision rule: if $(E, \text{high})$ and $(Q, \text{good})$ then $(P, \text{profit})$.

With every decision rule $\Phi \rightarrow \Psi$ we associate a conditional probability that $\Psi$ is true in $S$ given $\Phi$ is true in $S$ with the probability $\pi_S(\Phi)$, called the certainty factor of the decision rule

$$\pi_S(\Psi \mid \Phi) = \frac{\text{card}([\Phi \land \Psi]_S)}{\text{card}([\Phi]_S)},$$

where $[\Phi]_S$ denotes the set of all objects in $S$ having properties expressed by the formula $\Phi$.

Besides, we will also need the coverage factor of the decision rule

$$\pi_S(\Phi \mid \Psi) = \frac{\text{card}([\Phi \land \Psi]_S)}{\text{card}([\Psi]_S)},$$

which is the conditional probability that $\Phi$ is true in $S$ given $\Psi$ is true in $S$ with the probability $\pi_S(\Psi)$.

The certainty factor and coverage factor for the decision rules in Table 1 are given in Table 2.

<table>
<thead>
<tr>
<th>Store</th>
<th>$E$</th>
<th>$Q$</th>
<th>$L$</th>
<th>$P$</th>
<th>Cer.</th>
<th>Cov.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>high</td>
<td>good</td>
<td>no</td>
<td>profit</td>
<td>1</td>
<td>1/3</td>
</tr>
<tr>
<td>2</td>
<td>med.</td>
<td>good</td>
<td>no</td>
<td>loss</td>
<td>1/2</td>
<td>1/3</td>
</tr>
<tr>
<td>3</td>
<td>med.</td>
<td>good</td>
<td>no</td>
<td>profit</td>
<td>1/2</td>
<td>1/3</td>
</tr>
<tr>
<td>4</td>
<td>no</td>
<td>avg.</td>
<td>no</td>
<td>loss</td>
<td>1</td>
<td>1/3</td>
</tr>
<tr>
<td>5</td>
<td>med.</td>
<td>avg.</td>
<td>yes</td>
<td>loss</td>
<td>1</td>
<td>1/3</td>
</tr>
<tr>
<td>6</td>
<td>high</td>
<td>avg.</td>
<td>yes</td>
<td>profit</td>
<td>1</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Table 2. Certainty and coverage factors
In what follows we will distinguish two kinds of decision rules.

If \( \pi_S(\Psi|\Phi) = 1 \) then the decision rule \( \Phi \rightarrow \Psi \) is called certain. For example, the decision rule if \( (E, \text{high}) \) and \( (Q, \text{good}) \) and \( (L, \text{no}) \) then \( (P, \text{profit}) \) is certain.

If \( \pi_S(\Psi|\Phi) < 1 \) then the decision rule \( \Phi \rightarrow \Psi \) is called possible. For example, the decision rule if \( (E, \text{med.}) \) and \( (Q, \text{good}) \) and \( (L, \text{no}) \) then \( (P, \text{loss}) \) is possible.

Decision rules can be used to describe approximations. For example, the data set shown in Table 1 can be represented by the following minimal set of decision rules:

1) if \((E, \text{high})\) then \((P, \text{profit})\),
2) if \((E, \text{med.})\) and \((Q, \text{good})\) then \((P, \text{profit})\),
3) if \((E, \text{no})\) or \(((E, \text{med.})\) and \((Q, \text{avg.})\)) then \((P, \text{loss})\),
4) if \((E, \text{med.})\) and \((Q, \text{good})\) then \((P, \text{loss})\).

The rules 1) and 3) are certain decision rules and correspond to the lower approximations of sets of stores having profit and loss, respectively, whereas rules 2) and 3) are possible decision rules and correspond to the boundary regions of the above sets.

Certainty and coverage factors for these rules are given in Table 3

<table>
<thead>
<tr>
<th>rule</th>
<th>certainty</th>
<th>coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>1</td>
<td>2/3</td>
</tr>
<tr>
<td>2)</td>
<td>1/2</td>
<td>1/3</td>
</tr>
<tr>
<td>3)</td>
<td>1</td>
<td>2/3</td>
</tr>
<tr>
<td>4)</td>
<td>1/2</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Table 3. Certainty and coverage factors for the set (\(^\star\)) of decision rules

Let \( \{ \Phi_i \rightarrow \Psi \}_n \) be a set of decision rules such that:

all conditions \( \Phi_i \) are pairwise mutually exclusive, i.e., \( \| \Phi_i \land \Phi_j \|_S = \emptyset \), for any \( 1 \leq i, j \leq n, i \neq j \),

and

\[
\sum_{i=1}^{n} \pi_S(\Phi_i | \Psi) = 1
\]  

Then the following property holds:

\[
\pi_S(\Psi) = \sum_{i=1}^{n} \pi_S(\Psi | \Phi_i) \cdot \pi_S(\Phi_i).
\]  

(2)

For any decision rule \( \Phi_j \rightarrow \Psi \) the following is true:

\[
\pi_S(\Phi_j | \Psi) = \frac{\pi_S(\Psi | \Phi_j) \cdot \pi_S(\Phi_j)}{\sum_{i=1}^{n} \pi_S(\Psi | \Phi_i) \cdot \pi_S(\Phi_i)}.
\]  

(3)

The formula (2) is well known in probability calculus.
It can be easily seen that the relationship between the certainty factor and the coverage factor, expressed by the formula (3) is the Bayes’ rule. However, the meaning of Bayes’ rule in this case differs from that postulated in statistical inference. In statistical data analysis based on Bayes’ rule, we assume that prior probability about some parameters without knowledge about the data is given. The posterior probability is computed next, which tells us what can be said about prior probability in view of the data. In the rough set approach the meaning of Bayes’ rule is unlike. It reveals some relationships in the database, without referring to prior and posterior probabilities, and it can be used to reason about data in terms of approximate (rough) implications. Thus, the proposed approach can be seen as a new model for Bayes’ rule, and offers a new approach to data analysis.

The Bayes’ rule can be used to „inverse” the decision rules. With every decision rule if $\Phi$ then $\Psi$ we can associate an “inverse” decision rule if $\Psi$ then $\Phi$. For example, the set of inverse decision rules for the set of rules (*) is given below:

1) if $(P, \text{profit})$ then $(E, \text{high})$,
2) if $(P, \text{profit})$ then $(E, \text{med.})$ and $(Q, \text{good})$,
3) if $(P, \text{loss})$ then $(E, \text{no})$ or $(E, \text{med.})$ and $(Q, \text{avg.})$,
4) if $(P, \text{loss})$ then $(E, \text{med.})$ and $(Q, \text{good})$.

The inverse decision rule can be understood as an explanation of decisions in terms of conditions (i.e., giving reasons for decisions).

Certainty and coverage factors for inverse decision rules are given in Table 4.

<table>
<thead>
<tr>
<th>rule</th>
<th>certainty</th>
<th>coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>2/3</td>
<td>1</td>
</tr>
<tr>
<td>2)</td>
<td>1/3</td>
<td>1/2</td>
</tr>
<tr>
<td>3)</td>
<td>2/3</td>
<td>1</td>
</tr>
<tr>
<td>4)</td>
<td>1/3</td>
<td>1/2</td>
</tr>
</tbody>
</table>

Table 4. Certainty and coverage factors for the set (***) of inverse decision rules

The Bayes’ rule (3) and the formula (2) allow us to compute the probability of reasons. For example, for the decision rule if $(P, \text{profit})$ then $(E, \text{high})$ we have

$$
\pi((E, \text{high}) | (P, \text{profit})) = 2/3
$$

and

$$
\pi(P, \text{profit}) = 1/2
$$

hence

$$
\pi(E, \text{high}) = 2/3 \cdot 1/2 = 1/3.
$$

Thus, the probability that high empowerment of sales personnel gives profit is 1/3.
EXAMPLE OF APPLICATION

In this section we will discuss briefly the application of the rough set approach to the rotary clinker kiln control (Mrózek, 1989). Fig. 3 shows the simplified scheme of the kiln.

![Rotary clinker kiln](image)

The aim of the control is to mimic the behavior of the stoker of the kiln. To this end the control algorithm (set of control rules) has been generated from the analysis of the stoker behavior.

The stoker observes the burning zone of the kiln and identifies the state of the kiln by evaluating the following parameters, (condition attributes):

- $c_1 =$ burning zone temperature
- $c_2 =$ burning zone color
- $c_3 =$ clinker granulation in burning zone
- $c_4 =$ inside color of the kiln

Values of these parameters are given below:

Values of these parameters range as follows:

- $V_{c_1} = \{1,2,3,4\}$, where $1=[1380^\circ C-1420^\circ C]$, $2=[1421^\circ C-1440^\circ C]$, $3=[1441^\circ C-1480^\circ C]$, $4=[1481^\circ C-1500^\circ C]$
- $V_{c_2} = \{1,2,3,4,5\}$, where $1=$ scarlet, $2=$ dark pink, $3=$ bright pink, $4=$ decidedly bright pink, $5=$ rosy white
- $V_{c_3} = \{1,2,3,4\}$, where $1=$ fines, $2=$ fines with small lumps, $3=$ distinct granulation, $4=$ lumps
- $V_{c_4} = \{1,2,3\}$, where $1=$ distinct dark streaks, $2=$ indistinct dark streaks, $3=$ no dark streaks

Note that condition attribution are both quantitative (burning zone temperature) and qualitative (burning zone color, clinker granulation in the burning zone and inside color of the kiln).

After identification of the kiln state, determined by the condition attributes, the stoker using his knowledge and experience acts accordingly. His control decisions consist in setting values of the two following control parameters (decision attributes):

- $d_1 =$ kiln revolutions
- $d_2 =$ coal worm revolutions
Values of these parameters range as follows:

\[ V_{d_1} - \{1,2\}, \text{ where } 1=0,9[rpm], 2=1,22[rpm], \]

\[ V_{d_2} - \{1,2,3,4\}, \text{ where } 1=0[rpm], 2=15[rpm], 3=20[rpm], 4=40[rmp]. \]

In Table 5 (the decision table) control decisions of the stocker during one shift are given.

<table>
<thead>
<tr>
<th>Situation number</th>
<th>Condition attributes</th>
<th>Decision attributes</th>
<th>Situation number</th>
<th>Condition attributes</th>
<th>Decision attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>c₁ c₂ c₃ c₄</td>
<td>d₁ d₂</td>
<td></td>
<td>c₁ c₂ c₃ c₄</td>
<td>d₁ d₂</td>
</tr>
<tr>
<td>1</td>
<td>3 1 3 2</td>
<td>2 4</td>
<td>36</td>
<td>4 2 3 3</td>
<td>2 2</td>
</tr>
<tr>
<td>2</td>
<td>3 2 3 2</td>
<td>2 3</td>
<td>37</td>
<td>4 2 4 3</td>
<td>2 2</td>
</tr>
<tr>
<td>3</td>
<td>3 1 3 2</td>
<td>2 4</td>
<td>38</td>
<td>4 1 4 3</td>
<td>2 2</td>
</tr>
<tr>
<td>4</td>
<td>4 2 3 2</td>
<td>2 2</td>
<td>39</td>
<td>4 1 3 3</td>
<td>2 2</td>
</tr>
<tr>
<td>5</td>
<td>4 2 4 2</td>
<td>2 2</td>
<td>40</td>
<td>4 1 3 2</td>
<td>2 2</td>
</tr>
<tr>
<td>6</td>
<td>4 2 4 3</td>
<td>2 2</td>
<td>41</td>
<td>4 1 3 2</td>
<td>2 2</td>
</tr>
<tr>
<td>7</td>
<td>4 1 4 3</td>
<td>2 2</td>
<td>42</td>
<td>4 3 3 2</td>
<td>2 2</td>
</tr>
<tr>
<td>8</td>
<td>4 1 3 3</td>
<td>2 2</td>
<td>43</td>
<td>4 3 3 2</td>
<td>2 2</td>
</tr>
<tr>
<td>9</td>
<td>4 1 3 2</td>
<td>2 2</td>
<td>44</td>
<td>3 3 2 2</td>
<td>2 4</td>
</tr>
<tr>
<td>10</td>
<td>4 3 3 2</td>
<td>2 2</td>
<td>45</td>
<td>3 3 2 2</td>
<td>2 4</td>
</tr>
<tr>
<td>11</td>
<td>3 1 3 2</td>
<td>2 4</td>
<td>46</td>
<td>3 1 2 2</td>
<td>2 4</td>
</tr>
<tr>
<td>12</td>
<td>3 1 3 2</td>
<td>2 4</td>
<td>47</td>
<td>3 2 2 2</td>
<td>2 4</td>
</tr>
<tr>
<td>13</td>
<td>3 1 3 2</td>
<td>2 4</td>
<td>48</td>
<td>3 2 3 2</td>
<td>2 3</td>
</tr>
<tr>
<td>14</td>
<td>3 3 3 2</td>
<td>2 3</td>
<td>49</td>
<td>3 2 3 2</td>
<td>2 3</td>
</tr>
<tr>
<td>15</td>
<td>3 3 2 2</td>
<td>2 4</td>
<td>50</td>
<td>3 2 3 3</td>
<td>2 3</td>
</tr>
<tr>
<td>16</td>
<td>3 1 2 2</td>
<td>2 4</td>
<td>51</td>
<td>3 2 3 3</td>
<td>2 3</td>
</tr>
<tr>
<td>17</td>
<td>3 2 2 2</td>
<td>2 4</td>
<td>52</td>
<td>3 2 3 3</td>
<td>2 3</td>
</tr>
<tr>
<td>18</td>
<td>3 2 3 2</td>
<td>2 3</td>
<td>53</td>
<td>4 1 3 3</td>
<td>2 2</td>
</tr>
<tr>
<td>19</td>
<td>3 2 3 2</td>
<td>2 3</td>
<td>54</td>
<td>4 3 3 3</td>
<td>2 2</td>
</tr>
<tr>
<td>20</td>
<td>4 2 3 2</td>
<td>2 2</td>
<td>55</td>
<td>3 3 2 3</td>
<td>2 3</td>
</tr>
<tr>
<td>21</td>
<td>4 2 4 2</td>
<td>2 2</td>
<td>56</td>
<td>3 3 2 2</td>
<td>2 4</td>
</tr>
<tr>
<td>22</td>
<td>4 2 4 3</td>
<td>2 2</td>
<td>57</td>
<td>3 3 2 2</td>
<td>2 4</td>
</tr>
<tr>
<td>23</td>
<td>4 1 4 3</td>
<td>2 2</td>
<td>58</td>
<td>2 3 2 2</td>
<td>1 4</td>
</tr>
<tr>
<td>24</td>
<td>4 1 3 3</td>
<td>2 2</td>
<td>59</td>
<td>2 3 2 1</td>
<td>1 4</td>
</tr>
<tr>
<td>25</td>
<td>4 1 3 2</td>
<td>2 2</td>
<td>60</td>
<td>2 3 1 1</td>
<td>1 4</td>
</tr>
<tr>
<td>26</td>
<td>4 3 3 2</td>
<td>2 2</td>
<td>61</td>
<td>2 1 1 1</td>
<td>1 4</td>
</tr>
<tr>
<td>27</td>
<td>3 3 3 2</td>
<td>2 3</td>
<td>62</td>
<td>2 2 1 1</td>
<td>1 4</td>
</tr>
<tr>
<td>28</td>
<td>3 3 3 2</td>
<td>2 3</td>
<td>63</td>
<td>2 2 2 1</td>
<td>1 4</td>
</tr>
<tr>
<td>29</td>
<td>3 3 2 2</td>
<td>2 4</td>
<td>64</td>
<td>3 2 3 2</td>
<td>2 3</td>
</tr>
<tr>
<td>30</td>
<td>3 3 2 2</td>
<td>2 4</td>
<td>65</td>
<td>3 1 3 2</td>
<td>2 4</td>
</tr>
<tr>
<td>31</td>
<td>3 1 2 2</td>
<td>2 4</td>
<td>66</td>
<td>3 1 3 2</td>
<td>2 4</td>
</tr>
<tr>
<td>32</td>
<td>3 2 2 2</td>
<td>2 4</td>
<td>67</td>
<td>3 1 3 2</td>
<td>2 4</td>
</tr>
<tr>
<td>33</td>
<td>3 2 2 2</td>
<td>2 4</td>
<td>68</td>
<td>3 1 3 2</td>
<td>2 4</td>
</tr>
<tr>
<td>34</td>
<td>3 2 3 2</td>
<td>2 3</td>
<td>69</td>
<td>3 1 3 2</td>
<td>2 4</td>
</tr>
<tr>
<td>35</td>
<td>3 2 3 3</td>
<td>2 3</td>
<td>70</td>
<td>3 1 3 2</td>
<td>2 4</td>
</tr>
</tbody>
</table>

Table 5. Protocol of stocker decisions
Emploing the rough set technique, not shown here, one can obtain from Table 5 the following set of control rules:

1) \( (c_1, 3) \land ((c_4, 1) \lor (c_4, 2)) \Rightarrow (d_1, 2) \land (d_2, 4) \)
2) \( (c_4, 2) \Rightarrow (d_1, 1) \land (d_2, 4) \)
3) \( (c_3, 2) \land (c_4, 3) \Rightarrow (d_1, 2) \land (d_2, 3) \)
4) \( (c_3, 3) \Rightarrow (d_1, 2) \land (d_2, 2) \)

For details see Mrózek (1989).

The quality of control by the stoker and the rough control algorithm is revealed in Table 6.

<table>
<thead>
<tr>
<th>Name of parameter</th>
<th>Value of parameter during manual control</th>
<th>Value of parameter during rough controller control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigned temperature in burning zone (^{\circ}C)</td>
<td>1440</td>
<td>1440</td>
</tr>
<tr>
<td>Calculated mean temperature in burning zone (^{\circ}C)</td>
<td>1434</td>
<td>1435</td>
</tr>
<tr>
<td>Standard deviation of temperature in burning zone (^{\circ}C)</td>
<td>27.6</td>
<td>17.1</td>
</tr>
</tbody>
</table>

Table 6. Comparison of manual and automatic control

The fundamental parameter for the quality of the produced clinker is the temperature in of the burning zone. It is the same in both cases, i.e., by manual and automatic control. Hence both control methods perform equally well. However, there is an essential difference in the mean deviation of the temperature in the burning zone between the manual and automatic control: in the case of automatic control it is much less than in the case of manual control. This results in less coal consumption for heating of the kiln.

CONCLUSION

Rough set theory proved to be a very well suited candidate, beside fuzzy sets, neural networks and other soft computing methods, for intelligent industrial applications. Particularly challenging areas of applications of rough sets in industrial environment are material sciences, intelligent control, machine diagnosis and decision support.

Rough set approach has many advantageous features like, identifies relationships that would not be found using statistical methods, allows both qualitative and quantitative data and offers straightforward interpretation of obtained results.

Despite many successful applications of rough sets in industry there are still problems which require further research. In particular development of suitable, widely accessible software dedicated to industrial applications as well as microprocessors based on rough set theory are badly needed.
REFERENCES


15. Lin, T. Y.: Fuzzy controllers: an integrated approach based on fuzzy logic, rough sets, and evolutionary computing. In: T. Y. Lin and N. Cercone (eds.), Rough Sets and Data Min-


52. Szladow, A.: Datalogic/R: Mining the knowledge in databases. PC AI 7/1 (1993) 40-41


APPENDIX

A. ROUGH SETS FUNDAMENTALS


B. WHERE YOU CAN FIND MORE INFORMATION ABOUT ROUGH SETS

Electronic Bulletin of the Rough Set Community

Michael Hadjimichael (Editor)
Naval Research Laboratory, Monterey, USA
e-mail: hadjimic@nrlmry.navy.mil

Robert Golan (Asst. Editor)
Rough Knowledge Discovery Inc., Calgary, Canada
ftp: ftp.cs.uregina.ca:/pub/ebrcs

Rough Control Group Newsletter

Tosh Munakata (Chair)
Rough Set Bibliography

List of publications
http://papcio.ii.pw.edu.pl/roughbib.html

Bulletin of Informational Rough Set Society

S. Tsumoto, Y.Y. Yao, and M. Hadjimichael (Editors)
http://www.cs.uregina.ca/~roughset

C. ROUGH SET SOFTWARE

LERS - A Knowledge Discovery System
Jerzy W. Grzymala-Busse
University of Kansas, Lawrence, KS 66045, USA
e-mail: jerzy@eecs.ukans.edu

RSDM: Rough Sets Data Miner, A System to add Data Mining Capabilities to RDBMS
Maria C. Fernandez-Baizan, Ernestina Menasalvas Ruiz
e-mail: cfbaizan@fi.upm.es

ROSETTA
http://www.idi.ntnu.no/~aleks/rosetta.html
Andrzej Skowron
Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland
e-mail: skowron@mimuw.edu.pl

GROBIAN
http://www.infj.ulst.ac.uk/~ccc23/grobian/grobian.html
Ivo Düntsch, Günther Gediga
School of Information and Software Engineering,
University of Ulster, Newtownabbey, BT 37 0QB,
N.Ireland, e-mail: I.Duentsch@ulst.ac.uk
FB Psychologie Methodenlehre, Universität Osnabrück, 49069 Osnabrück, Germany, e-mail: ggediga@luce.psycho.Uni-Osnabrueck.DE

TRANCE: a Tool for Rough Data Analysis, Classification, and Clustering
Wojciech Kowalczyk
Vrije Universiteit Amsterdam De Boelelaan 1081A, 1081 HV Amsterdam, The Nether-
lands
e-mail:wojtew@cs.vu.nl

PROBROUGH - A System for Probabilistic Rough Classifiers Generation
Andrzej Lenarcik, Zdzislaw Piasta
Kielce University of Technology, Mathematics Department,
Al. 1000-lecia P.P. 5, 25-314 Kielce, Poland
e-mail: {zpiasta,lenarcik}@sabat.tu.kielce.pl

ROUGH FAMILY - Software Implementation of the Rough Set Theory
Roman Slowinski and Jerzy Stefanowski
Institute of Computing Science, Poznan University of Technology,
3A Piotrowo Street, 60-965 Poznan, Poland,
e-mail: Roman.Slowinski@cs.put.poznan.pl, Jerzy.Stefanowski@cs.put.poznan.pl

TAS: Tools for Analysis and Synthesis of Concurrent Processes using Rough Set Methods
Zbigniew Suraj
Institute of Mathematics, Pedagogical University, Rejtana 16A, 35-310 Rzeszow, Poland
e-mail: zsuraj@univ.rzeszow.pl

ROUGH FUZZY LAB - a System for Data Mining and Rough and Fuzzy Sets Based Classification
Roman W. Swiniarski
San Diego State University, San Diego, California 92182-7720, U.S.A.

PRIMEROSE
Shusaku Tsumoto
Medical Research Institute,
Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku Tokyo 113 Japan
e-mail: tsumoto@computer.org

KDD-R: Rough Sets-Based Data Mining System
Wojciech Ziarko
Computer Science Department, University of Regina, Regina
Saskatchewan, S4S-0A2, Canada
e-mail: ziarko@cs.uregina.ca