COENOSES 11(2): 55-67, 1996

FUZZY ROUGH SETS AS A TOOL FOR ECOLOGICAL DATA ANALYSIS

Ireneusz R. Moraczewski

Department of Plant Systematics and Geography, Warsaw University, Al. Ujazdowskie 4, PL-00-478 Warsaw, Poland.
E-mail: moraczew@bot.astrouw.edu.pl

Keywords: City zonation, Classification, Cluster analysis, Fuzzy set, Genetic algorithm, K-nearest neighbor algorithm, Partition, Rough set, Urban flora

Abstract: Rough set theory in its both crisp and fuzzy versions is presented. It is argued that the applicability of crisp rough sets is limited to some problem situations, whereas their fuzzy counterparts constitute a powerful tool for coping with complexities of real-world data. It is shown that the strength of a relationship between two sets of attributes C and D can be expressed by the quality of approximation of the similarity relation R_D by R_C. The goal of the analysis employing fuzzy rough sets is to find the smallest possible subset (or subsets) of condition attributes C^* for which the quality of approximation of the similarity relation R_D is the highest. The problem of combinatorial explosion involved in this task can be resolved by application of a genetic algorithm. Theoretical considerations are illustrated by a geobotanical example pertaining to the floristic description of anthropopressure zones in eight Polish cities.

Introduction

> *"It is therefore sufficient when speaking on such a subject and with such premises to indicate the truth roughly and in outline [...] for it is a part of an educated man to expect precision in any branch of science only in so far as the nature of the subject allows; it is surely just as absurd to admit probable reasoning from a mathematician as to ask a rhetorician for scientific demonstration."*
>
> Aristotle, Ethics, Book I.

An ecologist at work resembles a rhetorician rather than a mathematician. He/she deals with complex systems which, according to Zadeh's rule (Zadeh 1972), are subjected to approximate descriptions only. Uncertainty is an attribute of knowledge which an ecologist is faced with every day.

There are several distinct types of uncertainty. Each of them requires a different mathematical armamentarium for its appropriate description (Klir 1990). Probability theory captures the imperfection of knowledge brought about by uncertainty of event occurrence. Fuzzy set theory is capable of characterizing the type of uncertainty emerging from vagueness, ambiguity and imprecision of events, terms, relations, etc. Both formal tools are widely used nowadays.

This paper is devoted to a third kind of uncertainty that requires yet another way of conceptualization. Loosely speaking, the uncertainty in question is associated with an excessive refinement of an aspect of the domain compared with much coarser concepts used for its description. A group of objects different with respect to an attribute A can be indiscernible by an attribute B or even a set of attributes $B_1,...,B_k$. This can have important cognitive consequences. Consider a simple example provided in Fig. 1. Suppose each of the five figures is described by two binary attributes: *shape* (*circle* vs. *square*) and *color* (*white* vs. *black*). Should we define *shape* in terms of *color* we would be able to give a partial definition only:

color(white) → *shape(square).*

By applying this definition we may determine the shape of two objects out of five. The remaining three figures constitute the *uncertainty area* of the universe. None of them can be said to be black or white if their shape is the only item of information available. These objects are indiscernible by the attribute that we have at our disposal. Equivalence classes of the indiscernibility relation induced by the attribute *color* are too broad to permit an adequate definition of *shape*. This problem can be avoided if we enrich the definition with the size (*small* vs. *big*) of the figures. The addition of that attribute would result in splitting the equivalence class of previously indiscernible black figures into smaller units. This, in turn, would enable the formulation of an equivalence definition:

shape(circle) = *color(black)* and *size(big)*,

shape(square) = *color(white)* or *size(small)*.

A calculus of the indiscernible was introduced by Pawlak (1982, 1984, 1991) under the name of *rough set theory*. Recently, it has been applied successfully to various domains: medicine (Pawlak et al. 1986, Słowiński et al. 1988, Kandulski et al. 1992), pharmacology (Krysiński 1992), industrial processes control (Mrózek 1992), robotics (Ziarko 1992), logic (Nakamura 1992) and many other fields. In original formulation it exhibits, however, some

Figure 1. It is impossible to define the shape of figures merely on the basis of their color. The addition of the attribute *size* enables the formulation of an adequate definition. Further explanations in the text.

shortcomings which narrow its range of application. These are especially glaring in the case of such domains as ecology. The main limiting factor is an explicitly stated assumption that indiscernibility relation is a two-valued equivalence. Since in the field of ecology indiscernibility of objects (sample plots, stands, transects, communities, associations, niches, etc.) is a matter of degree, equivalence relations become vague, which in turn precludes crisp rough sets from serving as a means of data analysis.

Both types of knowledge imperfection i.e., granularity and vagueness, can be conjointly modeled. Different though interrelated approaches were suggested by Nakamura (1988), Dubois & Prade (1990, 1992) and Nanda & Majumdar (1992). These authors introduced the notion of a fuzzy rough set and studied its mathematical properties. However, they did not present fuzzy versions of many rough-set-theoretical concepts and techniques which are important from a practical viewpoint.

This paper presents the generalized (fuzzified) versions of the basic notions and methods of rough set theory. The approach presented will be semiformal in nature and kept in the spirit of the works of Dubois & Prade (1990, 1992). The utility of fuzzy rough sets will be illustrated by a geobotanical example pertaining to the floristic description of anthropopressure zones in eight Polish cities.

Fuzzy sets

The notion of a fuzzy set was introduced by L.A. Zadeh (1965). Fuzzy set theory has been developed extensively since the seventies. Although initially some mathematicians considered it defective and needless (see Zadeh 1990), it not only bore fruits in the field of theoretical thought (e.g., Kosko 1990), but also proved useful in numerous applications, among others in ecology (e.g., Allen & Starr 1982, Roberts 1986, 1989, Feoli & Zuccarello 1986, 1988, Moraczewski 1993a, 1993b). Cultural renderings and technological consequences of fuzzy set theory are vividly presented by McNeill & Freiberger (1992).

Given a universe X, a crisp subset S can be defined either by identifying the objects that are members of S or by specifying a characteristic function

Φ_S: $X \rightarrow \{0,1\}$.

A fuzzy subset A is derived by generalizing the characteristic function to a membership function

μ_A: $X \rightarrow [0,1]$.

The following basic operations on fuzzy sets will be used in the sequel:

complement: $\mu_{\overline{A}}(x) = 1 - \mu_A(x)$, $\forall x \in X$,

union: $\mu_{A \cup B}(x) = MAX(\mu_A(x), \mu_B(x))$, $\forall x \in X$,

cardinality: $card(A) = \sum_{i=1}^{N} \mu_A(x_i)$,

where the N elements of X constitute the universe of discourse.

Rough sets

Before jumping into rough set theory with both feet it will be useful to outline the main idea of the theory and to present the intuitive distinction between fuzziness and roughness.

Let X be a universe of discourse and R a reflexive, symmetrical, and transitive indiscernibility relation inducing a set of mutually exclusive and non-empty equivalence classes. Let us assume that a proper subset S of the universe is given. The main question addressed by Pawlak (1982, 1984, 1991) is: how to represent set S by means of coarsening (the equivalence classes) of X?

Set S can be characterized with the use of two sets: the lower and the upper approximations (Fig. 2). The lower approximation is a collection of equivalence classes which are included in S. The upper approximation is a set of those equivalence classes whose intersection with S is non-empty. When the approximations do not match, S cannot be described perfectly. The approximations constitute then a rough description of S.

Roughness and vagueness are distinct types of uncertainty and can exist separately. Let us take a look at Fig. 3 showing the difference. Both letters on the right-hand side are an imperfect reflection of original letters on the left-hand side, but the distortions have different reasons. Imperfection in the upper row is caused by an increase in the pixel size (a decrease in resolution), whereas that in the lower one by a

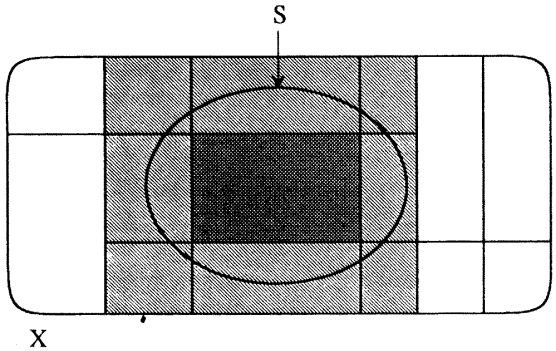


Figure 2. A reflexive, symmetrical and transitive relation R on a universal set X induces equivalence classes which are presented as mutually exclusive areas. A set $S \subset X$ can roughly be characterized by means of the lower (one-element set containing the central rectangle) and upper approximation (the collection of rectangles filled with a hatch pattern).

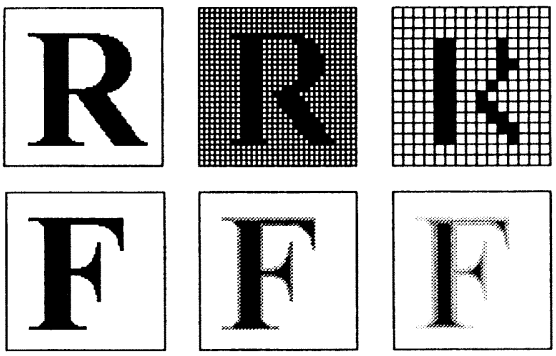


Figure 3. The imperfection of image can result from the roughness (upper row) or fuzziness (lower row). Further explanations in the text.

decrease in sharpness (contrast) at the letter's fringe. Of course, both phenomena can occur simultaneously and they usually do. The reality we have to deal with often resembles a pattern displayed on a screen divided into granules and allowing many shades of gray.

Fuzzy rough sets

Knowledge representation system

The knowledge of a domain can be represented as a 4-tuple referred to as information system (*IS*) (Pawlak 1991):

$IS = (X,A,V,f)$,

where $X = \{x_1,...,x_i,...,x_N\}$ is a finite set of objects (observations, events, samples, communities, associations, etc.) called the universe of discourse, A is a finite set of attributes (variables, characters, features, dimensions, species), and V is a sum of attribute domains, i.e.

$$V = \bigcup_{a \in A} V_a;$$

f is a mapping assigning a value from V to each ordered pair (object, attribute)

$f: X \times A \rightarrow V$.

Our knowledge relates to a finite set X which is usually a part of a potentially infinite set U. This fact should be borne in mind when applying the rules obtained for elements of X to the objects from U-X.

Indiscernibility and similarity relations

The basic notion of the rough set theory in its original formulation (Pawlak 1982, 1984, 1991) is the indiscernibility relation R. Two objects x, y belonging to X are indiscernible by a collection of attributes $A' \subseteq A$ if and only if

$f(x,a) = f(y,a)$, $\forall\, a \in A'$,

where $f(x,a)$ and $f(y,a)$ denote the value of the attribute a taken by the objects x and y, respectively. Since R_i is reflexive, symmetrical and transitive, it induces a classification of objects from X, i.e. a collection of mutually exclusive and non-empty equivalence classes: $X/R = \{[x_i]_R\} = \{X_1, ..., X_i, ..., X_n\}$. This set is called a coarsening of X; it is the starting point for defining the key concepts of the theory at issue.

The usefulness of the theory depends, however, on whether or not an indiscernibility relation is an appropriate model of the reality analyzed. Of course, such a relation can be created for every information system, but in many cases it proves useless. This applies to quantitative data (measurements in ratio or interval scales). For such data, R in most cases will be a degenerate relation containing only the identity pairs (x,x).

The second limitation of the applicability of crisp indiscernibility relations is not an inherent feature of any kind of data; it can be manifested even in a boolean data set. It is connected with the nature of objects. To be more specific, consider ecological systems. The latter are often described by tens, hundreds or even thousands of attributes. As phenomena that come spontaneously into being they exhibit enormous variability in space and time. That is why any two objects treated by ecologists as "basically the same" (e.g., two adjacent stands of *Tilio-Carpinetum*) are not usually identical in terms of species composition. This fact, discovered many decades ago, has found its expression in research work: ecologists used to speak about the *similarity* of objects instead of ascertaining their identity.

In order to expand the applicability of rough sets we must replace the concept of indiscernibility with a similarity relation. Consequently, the crisp partition must give way to its fuzzy counterpart.

The concept of a similarity relation was introduced by Zadeh in 1971. A similarity relation is a fuzzy subset of the cartesian product $X \times X$ such that

1) $\mu_R(x,x) = 1$, $\forall x \in X$ (reflexivity),

2) $\mu_R(x,y) = \mu_R(y,x)$, $\forall\, x,y \in X$ (symmetry),

3) $\mu_R(x,z) \geq \underset{y \in X}{MAX} (\mu_R(x,y) * \mu_R(y,z))$, $\forall\, x, z \in X$ (transitivity),

where * stands for some operator satisfying $a * b \leq MIN\,(a, b)$. It is usually assumed that $* = MIN$. For other types of admissible transitivity axioms see Bezdek & Harris (1978). It can be demonstrated easily (Zadeh 1971, Dunn 1974) that a similarity relation with the *MAX-MIN* transitivity implies (and is implied by) an ultrametrics.

A fuzzy equivalence class $[x_i]_R$, denoted also by Φ_i, we define after Zadeh (1971) as follows:

$\mu_{[x_i]R}(y) = \mu_R(x_i,y)$, $\forall\, y \in X$.

In ecology the procedure of creating similarity relations (or, rather, ultrametrics corresponding to them) consists of two stages. It includes, first, the computation of a reflexive and symmetrical distance matrix $\mathbf{D} = [d_{xy}]$. The next stage of the procedure consists in the transformation of $\mathbf{D}$ into an ultrametrics $\mathbf{D}^U = [d^U_{xy}]$ using a clustering technique. Since the values of $\mathbf{D}^U$ need not (and often do not) fall into the [0,1] interval, and the membership grades of a fuzzy relation R dual to $\mathbf{D}^U$ should meet this requirement, we assume that R,

unless it is explicitly given, is defined as a complement of a standardized version of $\mathbf{D}^U$:

$$\mu_R(x,y) = 1 - \delta_{xy}, \quad \forall\, x,y \in X, \tag{1}$$

where

$$\delta_{xy} = \frac{d^U_{xy} - \underset{i\neq j}{MIN}\, d^U_{ij}}{\underset{i\neq j}{MAX}\, d^U_{ij} - \underset{i\neq j}{MIN}\, d^U_{ij}} \tag{2}$$

It is clear from (1) and (2) that the relation R will contain at least one pair (x,y), $x\neq y$, such that $\mu_R(x,y) = 1$. Moreover, the objects which according to **D** are the most distant ones will become totaly different in terms of R.

Although nothing particular about the distance measure as well as the clustering technique is assumed here, their selection usually affects the results and therefore should be made with extreme caution.

Following Dubois & Prade (1990, 1992) we now introduce an auxiliary notion. If $\Phi_i \in X/R$, then its extension (being a crisp set) is defined by

$\omega(\Phi_i) = \{x \mid$ 'Φ_i' is the name of the (fuzzy) equivalence class $[x]_R\}$.

An extension of a fuzzy set A in X/R, denoted by $\omega(A)$, we can define as:

$$\mu_{\omega(a)}(x) = \mu_A(\Phi_i), \text{ if } x \in \omega(\Phi_i), \forall\, x \in X.$$

Approximation of a set

Now we proceed to present the key notions of rough set theory in its orthodox (crisp) as well as fuzzified versions.

Let S be a crisp subset of an universal set X, $S \subset X$, and $X/R = \{X_1,...,X_i,... X_n\}$ a coarsening of X. The lower and upper approximations, $R_*(S)$ and $R^*(S)$, respectively, are defined as (Dubois & Prade 1990,1992):

$$R_*(S) = \{\, X_i \mid X_i \subseteq S, \forall\, X_i \in X/R\,\}\ , \tag{3}$$

$$R^*(S) = \{\, X_i \mid X_i \cap S \neq \emptyset, \ \forall\, X_i \in X/R\,\}. \tag{4}$$

Both definitions can be rewritten as follows:

$$\Phi_{R_*(S)}(X_i) = \underset{x\in X}{MIN}\, [\Phi_{X_i}(x) \rightarrow \Phi_S(x)], \quad \forall\, X_i \in X/R, \tag{5}$$

$$\Phi_{R^*(S)}(X_i) = \underset{x\in X}{MAX}\, [MIN(\, \Phi_{X_i}(x), \Phi_S(x)\,)], \quad \forall\, X_i \in X/R, \tag{6}$$

where $p\rightarrow q$ denotes the conventional implication:

$$p\rightarrow q \begin{cases} 0, & \text{if } p=1 \text{ and } q=0, \\ 1, & \text{otherwise} \end{cases} \tag{7}$$

A pair $(R_*(S), R^*(S))$ is referred to as (crisp) roughh set.

Now assume that the universal set X is equipped with a fuzzy partition $X/R = \{\Phi_1, \Phi_i, ...,\Phi_n\}$, and a fuzzy set $F \subset X$ is to be described by means of fuzzy equivalence classes of R. Then, the characteristic-function definitions (5)-(6) can be rewritten as follows (Dubois & Prade 1990,1992):

$$\mu_{R_*(F)}(\Phi_i) = \underset{x\in X}{MIN}\, [\mu_{\Phi_i}(x) \Rightarrow \mu_F(x)], \ \forall\, \Phi_i \in X/R\} \tag{8}$$

$$\mu_{R^*(F)}(\Phi_i) = \underset{x\in X}{MAX}\, [MIN(\mu_{\Phi_i}(x), \mu_F(x))], \ \forall\, \Phi_i \in X/R\}, \tag{9}$$

where $p\Rightarrow q$ is a fuzzy implication. A pair $(R_*(F), R^*(F))$ can be called a fuzzy rough set.

There are several definitions of generalized implications. Bouchon-Meunier (1992) listed and discussed the most frequently used ones. For the purpose of this work, the Kleene-Dienes implication has been employed:

$$p\Rightarrow q = MAX\,(1 - p,\, q\,). \tag{10}$$

In the crisp case, when p and q take values from the set $\{0,1\}$, *(10) coincides with the definition of classical implication* (7).

Note that (8)-(9) make sense even when F is a crisp set. This situation is observed in the example discussed below. A graphic illustration of determining the membership value of Φ_i in the lower approximation of a crisp set S is given in Fig. 4.

For the sake of clarity let us add that when both F and Φ_i, for $i = 1, ..., n$, are crisp sets, then (8)-(9) coincide with (5)-(6) and (3)-(4).

Among the concepts defined, the lower approximation is of greatest importance in practice. It enables the determination of degree of dependency between similarity relations induced by various collections of attributes. Here is an important property of the lower approximation. If Φ_i is a properly fuzzy equivalence class, then the degree of membership of Φ_i in its own lower approximation is less than 1, and *vice versa*:

$$(\exists\, x \in X:\ \mu_{\Phi_i}(x) < 1) \leftrightarrow \mu_{R_*(\Phi_i)}(\Phi_i) < 1 \tag{11}$$

Indeed, from the definition of the lower approximation we have

$$\begin{aligned} \mu_{R_*(\Phi_i)}(\Phi_i) &= \underset{x\in X}{MIN}\, [\mu_{\Phi_i}(x) \Rightarrow \mu_{\Phi_i}(x)], \\ &= \underset{x\in X}{MIN}\, [MAX(\mu_{\overline{\Phi}_i}(x), \mu_{\Phi_i}(x))], \\ &= \underset{x\in X}{MIN}\, \mu_{\overline{\Phi}_i \cup \Phi_i}(x). \end{aligned} \tag{12}$$

The left-hand side of (12) equals 1 if and only if $\overline{\Phi}_i \cup \Phi_i = X$, that is, when Φ_i is a crisp set, which contradicts the assumption.

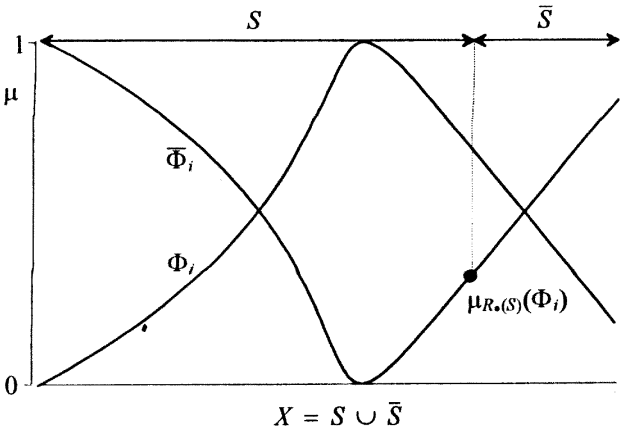


Figure 4. Determining the degree of membership of Φ_i in the lower approximation of a crisp set S.

Moreover, if $\mu_{R_*(\Phi_i)}(\Phi_i) < 1$, then $\mu_{R_*(\Phi_j)}(\Phi_i) < 1$, $\forall\ \Phi_j \in X/R$, since there always will exist such an x, that $\mu_{\Phi_i}(x) = 1$ (and $\mu_{\overline{\Phi}_i}(x) = 0$) and simultaneously $\mu_{\Phi_j}(x) < 1$ (because $\Phi_i \neq \Phi_j$). Thus,

if $\mu_{R_*(\Phi_i)}(\Phi_i) < 1$, then $\mu_{\omega(R_*(\Phi_j))}(x_i) < 1, \forall\ \Phi_j \in X/R$. (13)

In the next section we will refer to this property.

Approximation of a fuzzy partition

The attribute set A of an information system IS is not usually a homogenous one. An expert in the field is able, as a rule, to distinguish two mutually exclusive collections of attributes C, $D \subset A$, $C \cap D = \emptyset$. There are many possible interpretations of these classes, their reciprocal relationships and the research goals associated with them. We mention here the basic ones only:

- C covers the attributes (variables) being potential causes of the effects described by the attributes from D; a causal relation between the attributes is searched for;
- C is a set of condition attributes and D comprises the decision attributes; adequate and the smallest possible decision rules are to be found;
- C is a collection of attributes which can have a diagnostic value for some classes defined by the attributes belonging to D; an effective small subset of indicators is to be determined.

Given a distance matrix and a clustering technique, both attribute classes C and D induce fuzzy equivalence classes $X/R_C = \{\Gamma_1,...,\Gamma_i,...,\Gamma_n\}$ and $X/R_D = \{\Delta_1,...,\Delta_j,...,\Delta_m\}$, respectively. In order to express the dependency between these partitions we will introduce a measure called the quality of approximation of R_D by R_C. This measure can be interpreted as a strength of a causal nexus, or a degree to which the decisions (resp. diagnoses) described by the attributes from D are determined by the condition (resp. diagnostic) attributes.

We shall first be concerned with the requirement that has to be imposed on the characteristic of two objects from X, if the relationship between two sets of attributes C and D is to be interpreted in one of the three ways mentioned above. Let us begin with the crisp case. If x and y are indiscernible by C, they must be indiscernible by D, too. The basic idea of that requirement is very similar to Mill's method of concomitant variations. It rests on the assumption that as long as a cause does not change, the phenomenon which is the effect thereof does not change either. Otherwise we would arrive at an undesirable situation: identical causes would have different (mutually exclusive) effects. This also applies, *mutatis mutandis*, to the two remaining interpretations. It can easily be seen that all the elements belonging to the extensions of lower approximations of the equivalence classes from X/R_D satisfy the condition specified above. A situation when x and y are indiscernible by D but different with respect to the attributes from C is quite acceptable: an effect can be due to many causes. Analogously, various motives induce us to make a decision, and many symptoms can indicate a diagnosis.

In the fuzzified case this requirement becomes weaker. We demand that x and y close to each other under R_C, are similar with respect to R_D, too. The opposite need not (but it may) hold. Loosely speaking, we expect similar effects from similar causes. Elements of X belonging, with a certain degree of membership, to the extensions of lower approximations of Δ_j meet the requirement in question to this degree.

The sum over all the extensions of lower approximations of (fuzzy) equivalence classes Δ_j is called a positive region of R_D and denoted by $POS_{R_C}(R_D)$ or, for simplicity, $POS_C(D)$ (cf. Pawlak 1991):

$$POS_C(D) = \bigcup_{\Delta_j \in X/R_D} \omega(R_C{}_*(\Delta_j)).$$

Now, by referring to the concept of a positive region we can define the quality of approximation of R_D by R_C (also called the degree of dependency between R_C and R_D), in symbolical notation $\gamma_{R_C}(R_D)$ or, for simplicity, $\gamma_C(D)$ (cf. Pawlak 1991):

$$\gamma_C(D) = \frac{card(POS_C(D))}{card(X)}, \tag{14}$$

where $card(.)$ stands for the cardinality. The quality of approximation ranges over the [0,1] interval. The closer this coefficient approaches 1, the better the approximation of R_D by R_C.

We will demonstrate now the main difference between the quality of approximation defined by (14) and the so-called coefficients of cophenetic correlation (see, e.g., Jardine & Sibson 1971). If R_D is a properly fuzzy relation, and fuzzy implication involved in (8) is defined by (10), then (13) entails that

$$\gamma_D(D) < 1. \tag{15}$$

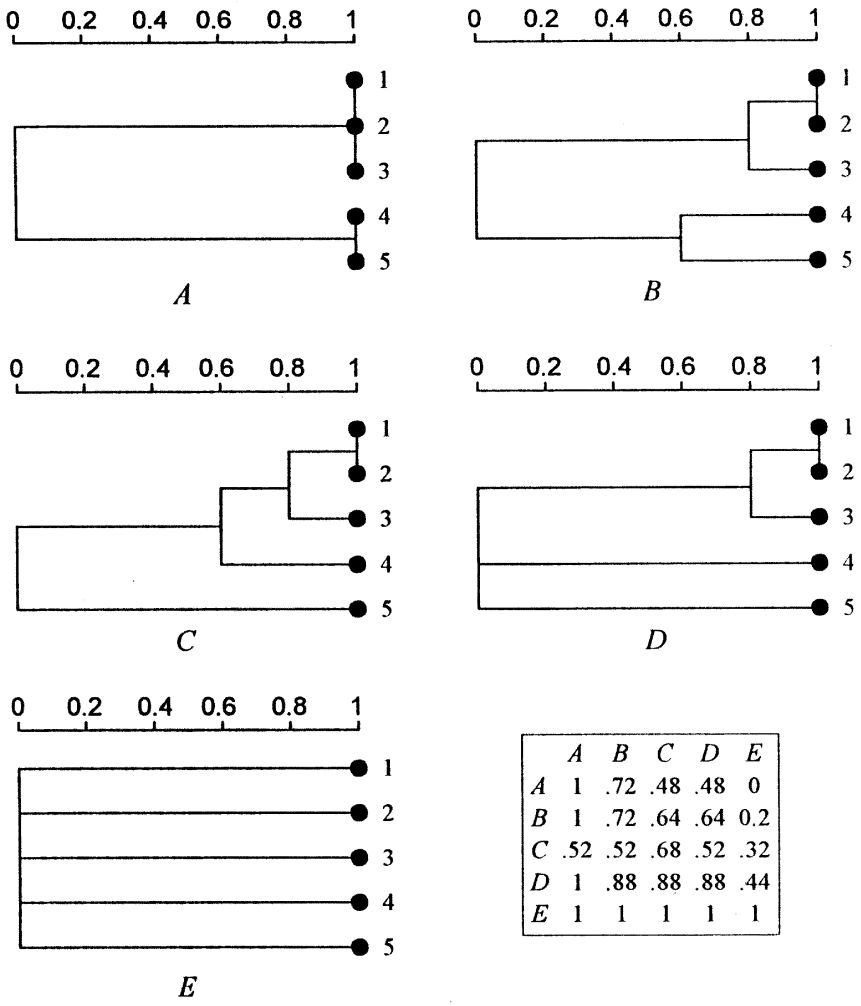
Hence, the quality of approximation does not exhibit the properties of reflexivity and symmetry (the value of $\gamma_C(D)$ as a rule does not equal to that of $\gamma_D(C)$). This measure is not a resemblance coefficient. Pawlak (1991) suggested that it can be interpreted as "an ability to classify objects". The property (15), which could be viewed as a paradox if the γ coefficient was considered a similarity measure, can be treated as a numerical reflection of the difficulty in classifying the objects: each of them belongs simultaneously to many classes.

Also note that

$$\gamma_I(D) = 1,$$

where I is a crisp identity relation, $I = \{(x,x), \forall\ x \in X\}$. This means that a relation in which each object is totaly different is perfectly consistent with any other relation.

For the sake of illustration of how the γ coefficient works, several exemplary fuzzy partitions (dendrograms) together with the degrees of dependency for each ordered pair are presented in Fig. 5.



	A	B	C	D	E
A	1	.72	.48	.48	0
B	1	.72	.64	.64	0.2
C	.52	.52	.68	.52	.32
D	1	.88	.88	.88	.44
E	1	1	1	1	1

Figure 5. Five artificial dendrograms (fuzzy partitions) *A-E* and values of the γ coefficient for each of the ordered pairs. Note that dendrograms *A* and *E* correspond to crisp partitions {{1,2,3},{4,5}} and {{1},{2},{3},{4},{5}}, respectively. $\gamma_A(B) = 0.72$, $\gamma_B(A) = 1.0$, etc.

Having defined the quality of approximation we can now proceed to discuss the problem of attribute selection.

Selection of attributes

Let a simple property be our starting point:

If R_C and R_D are crisp indiscernibility relations, then

$$\gamma_{C-\{a\}}(D) \leq \gamma_C(D), \tag{16}$$

where $C - \{a\}$ is the set C from which an attribute a has been removed. In order to prove (16) it is sufficient to demonstrate that

$$R_{C-\{a\}} \supseteq R_C,$$

hence

$$POS_{C-\{a\}}(D) \subseteq POS_C(D),$$

thus

$$card(POS_{C-\{a\}}(D)) \leq card(POS_C(D)).$$

The removal of an attribute from set C can make two different objects indiscernible, but never the other way round. Therefore, for high cardinalities of C the crisp indiscernibility relation is very likely to contain identity pairs only, which suffices for the γ coefficient to attain the maximum value of 1.

A change in the quality of approximation of R_D resulting from the removal of an attribute a from set C expresses the relative contribution of a to the entire dependency between R_C and R_D and is referred to as significance factor of the attribute a (Pawlak 1991):

$$SGF(a,C,D) = \frac{\gamma_C(D) - \gamma_{C-\{a\}}(D)}{\gamma_C(D)}$$

In the crisp case, the significance factor takes values from the [0,1] interval.

Let $\mathcal{C}$ be a set of all the subsets of A-D:

$$\mathcal{C} = \{C \mid C \subseteq A - D\}.$$

The main goal of the analysis employing crisp rough sets is to determine the smallest (in the sense of cardinality) subset $C^* \in \mathcal{C}$ called a reduct that

$$\gamma_{C^*}(D) = \gamma_{A-D}(D). \tag{17}$$

In reply to a simple question why we insist on finding the smallest sets fulfilling the above condition, Pawlak (1991) pointed out that it is due to economy and simplicity.

The problem of finding reducts is NP-hard (Ziarko 1991). In practice this means that any exhaustive search for reducts is not feasible even for low cardinalities of C. When C is not too big, many deterministic methods of search for small subsets (not necessarily the smallest ones) can be applied (Pawlak 1984, Tanaka et al. 1992, Orłowska & Orłowski 1992, Słowiński & Stefanowski 1992). Nevertheless, these methods are useless when the search becomes ensnared in a combinatorial explosion. For big sets of attributes it is necessary to use indeterministic (approximate) algorithms. In the work of Moraczewski et al. (in press) the search for subsets fulfilling (17) was performed with the use of a genetic algorithm.

The monotonicity property (16) is not satisfied if R_C and R_D are similarity relations. It can easily be demonstrated (by an appropriate example) that the removal of an attribute from set C can result in an increase in the quality of approximation; thus, $SGF(a,C,D)$ can be negative. Therefore, as opposed to the crisp case, we must perform a *selection* rather than a reduction of attributes. Hence, the main goal of the analysis with the use of fuzzy rough sets is to determine such a $C^* \in C$ for which $\gamma_{C^*}(D)$ attains its maximum value.

We present below an indeterministic technique for finding subsets of C that give the highest (or adequately high) quality of approximation.

It is good to start the search with one-element subsets of A-D, and then go on with subsets of increasing cardinality. This procedure ensures that once a subset yielding acceptable quality of approximation is found, it is relatively small at the same time. For low values of k the search can assume the form of a simple enumeration but finding greater subsets (starting with $k = 2$ for the dataset presented below) requires the application of approximate techniques, such as a genetic algorithm (GA).

Genetic algorithms, developed by Holland (1975), are optimization procedures that try to mimic the evolutionary process. Every GA starts with a random population of strings (chromosomes), each one representing a solution to the given problem. These solutions are improved in the course of "evolution" by means of three operators: mutation, crossover and duplication (see Goldberg 1989). The representation of solutions as well as the exploration operators used in this work borrow heavily from Lucasius et al. (1993). As the description of the GA implementation clearly goes beyond the scope of this paper, we will confine ourselves to the very basic explanations only. Each chromosome is a k-element subset of attributes (or, more precisely, a subset of numbers representing them). Mutation is a random alteration of one element in the subset. Crossover operator applies to a pair of chromosomes (subsets) and yields two subsets containing elements randomly chosen from the parental strings. The quality of approximation (14) can serve directly as the fitness function.

Decision rules

As already said in one of the first sections of this work, set X is usually a subset of a broader set U. The knowledge about elements of X drawn out from IS and presented in a form of decision rules can be applied to reasoning about objects from U-X.

Let $SI=(X,C\cup D,V,f)$ be an information system in which $C\cap D = \emptyset$. *Collections of attributes* C and D induce, according to (1), two sets of fuzzy equivalence classes $X/R_C = \{\Gamma_1,...,\Gamma_i,...,\Gamma_n\}$ and $X/R_D = \{\Delta_1,...,\Delta_j,...,\Delta_m\}$, respectively. An ordered couple $\rho_{ij} = (\Gamma_i,\Delta_j)$, $i = 1,...,n$, $j = 1,...,m$, is called a decision rule (Pawlak 1991). The set of all the decision rules will be denoted by P. Decision rules differ in their consistency. In the crisp case (when R_C and R_D are indiscernibility relations) a rule ρ_{ij} is consistent iff

$$\Gamma_i \subseteq \Delta_j, \tag{18}$$

otherwise it is inconsistent. In this case a collection of consistent rules P_{CNST} is a crisp subset of P. An inconsistent rule is useless (because indiscernible causes have different, mutually exclusive effects, etc.).

The sharp division into useful and useless rules becomes blurred in the case of fuzzy equivalence classes. The formula (18) rewritten in terms of membership functions takes the following form:

$$\mu_{P_{CNST}}(\rho_{ij}) = \underset{x\in X}{MIN}\,(\mu_{\Gamma_i}(x)\Rightarrow\mu_{\Delta_j}(x)),\ \forall\, x \in X,$$

where $p \Rightarrow q$ is a fuzzy implication defined by (10).

If both R_C and R_D are indiscernibility relations, a rule ρ_{ij} is interpreted as a classical conditional 'if ... then ...'. The antecedent (resp. consequent) of the conditional is a conjunction of descriptors (attributes equipped with a value) of the class Γ_i (resp. Δ_j).

Classical interpretation of decision rules, although very elegant from the logical viewpoint, has some serious disadvantages. We will point them out.

Suppose, for instance, that set X consists of 10 elements and C is also a 10-element set of binary attributes. Then set P will contain up to 10 consistent decision rules. This means that we will have some knowledge of merely ten combinations of values of the condition attributes, which covers only 1% of the whole space of possible combinations. If we now encountered a new combination which does not match any of the antecedents, the conventional form of our knowledge would prohibit us from performing any reasoning.

Another difficulty is also associated with the logical form of the antecedent. Since the latter is a conjunction, it is sufficient for it to be false that an examined object u differs from an x of X in one attribute (out of, say, a thousand). To alleviate this problem some methods of eliminating superfluous, irrelevant descriptors (Pawlak 1991, Grzymała-Busse 1992) are proposed. These techniques make the conjunctions more succinct, thereby scaling the problem down.

These problems can be best avoided by employing a generalized technique for utilizing decision rules. One way is building a fuzzy inference system based on a collection of fuzzy conditionals of the form 'if Γ_i then Δ_j' (Tanaka et al.

1992; see also Moraczewski 1993b). Another possibility is to make use of fuzzy *K*-nearest neighbor algorithm (fuzzy *K*-NN in short; Keller et al. 1985). This simple yet effective procedure will be presented in the sequel.

There are three assumptions underlying the fuzzy *K*-NN algorithm:

1) consistency (usefulness) of the rules is a matter of degree,

2) for a rule r_{ij} to be fired (applied) a conjunction of requirements need not be completely fulfilled; some degree of activation may be sufficient,

3) many rules can be fired at the same time.

The chief point of the method is to apply those rules, whose antecedents treated as character vectors are the closest to the vector corresponding to an examined object u. As a result of the algorithm's performance we get a set of m values $COMP(u,\Delta_j)$, $j = 1,...,m$ belonging to the [0,1] interval. For terminological reasons these values will be called degrees of compatibility (rather than membership) of an element u with the class Δ_j. Further proceeding depends on the goal of analysis. We are able, for example, to identify such a class for which the degree of compatibility is the highest, and then interpret it as the effect searched for, the decision recommended or the appropriate diagnosis - depending on the problem situation. This mode of analysis has been applied to exemplary computations reported in the next section.

Let $\mathbf{x}_i$ be a vector associated with the extension of fuzzy equivalence class ω ($[x_i]_{R_C}$):

$\mathbf{x}_i = [f(x_i,c_1),...,f(x_i,c_k),...,f(x_i,c_l)]$,

where $l = card(C)$. An analogous vector for the object u will be denoted by $\mathbf{u}$. The fuzzy version of *K*-nearest neighbor algorithm is summarized below.

```
Set K;  /* 1 ≤ K ≤ n */
s = 1;
for(i = 1; i ≤ n; i++)
    Compute d_xiU (i.e., a distance from x_i to u);
Standardize all the distances to the [0,1] interval;
for(i = 1; i ≤ n; i++) /* finding nearest neighbors */
{
    if(s ≤ K)
    {
        Include x_i in the set of nearest neighbors;
        s++;
    }
    else
        if(x_i is closer to u than any previous nearest
            neighbor)
        {
            Delete farthest in the set of nearest neighbors;
            Include x_i in the set of nearest neighbors;
        }
}
for(j = 1; j ≤ m; j++) /* determining compatibility of u with
        the decision classes */
    Compute COMP(u, Δ_j) using (19);
```

$$COMP(u,\Delta_j) = \begin{cases} 0 & \text{if } \sum_{r=1}^{K}(1-\delta_{x_r u}) = 0 \\ \dfrac{\sum_{r=1}^{K} \mu_{P_{CNST}}(\rho_{r_j}) * (1-\delta_{x_r u}) * \mu_{\Delta_j}(x_r)}{\sum_{r=1}^{K}(1-\delta_{x_r u})} & \text{otherwise} \end{cases} \tag{19}$$

where $\delta_{x_r u}$ is a standardized distance between $\mathbf{x}_r$, $r = 1,...,K$, and $\mathbf{u}$.

K is a small natural number. Too high as well as too low values of *K* tend to lower the efficiency of prediction (Tadeusiewicz & Flasiński 1991).

Let us emphasize that in the crisp case, when the distances between vectors and the degree of consistency take their values from the set {0,1}, the presented algorithm coincides with the conventional (i.e., involving 'if... then...' rules) procedure of reasoning.

A floristic example

Studies on urban floras

The investigations of urban floras have a long history. Floristic data for some European cities were collected as early as in the 17th century, e.g., Bauhin's (1622) data for Basel, Bernhardi's (1652) data for Warsaw. Since the first half of the 19th century more botanists have turned their attention to the flora of towns (e.g., Klett & Richter 1830, Hacker 1844, Garcke 1848, etc.). Floristic research in towns has been intensified for the past 30 years. It is estimated that the flora of about 50 larger European cities has already been examined (cf. Klotz 1989, 1990, Brandes 1989, Pyšek 1993, 1994 and the literature cited there).

Only recently floristic studies of urban areas have been conducted in a more systematic way, i.e. within a grid (the area of the grid squares is usually 0.5 x 0.5, 1.0 x 1.0 or 2.0 x 2.0 km). This method has already been applied to the flora of London (Burton 1983), Münster (Gödde & Wittig 1983), Köln (Kunick 1984), Warsaw (Sudnik-Wójcikowska 1987), Zürich (Landolt 1991), Poznań (Jackowiak 1993), and that of many other cities.

Analysis of the floristic records provided an insight into the ecology of expansive and declining species, and contributed to better understanding of changes of flora affected by increasing human impact. However, when carrying out research into the relationships between taxa and habitat factors, a fact must be borne in mind. Relatively big spacing of the grid overlay implies the heterogeneity of the squares, which in turn entails that the values of parameters ascribed

to them are only crude estimates. The same applies to the flora itself: a more or less continuous distribution of each species must be expressed by a pattern consisting of relatively big granules. That is why each statement being a crowning achievement of the florists must bear the seal of rough knowledge.

Anthropopressure as a complex factor influencing the urban flora

Studies concerning the human influence on the biodiversity in cities are of particular importance. A pre-condition of such a research is defining the anthropopressure.

The problem of city zonation based on the intensity of human impact is not a new one. More or less successful attempts were made in the 70's. (cf. Kunick 1974, Rapoport et al. 1983, Wittig et al. 1985, Jackowiak 1993). When distinguishing these zones the most difficult problem is the necessity to combine simultaneously a great number of factors. Only some of them can be gauged directly. To make things worse let us recall that each grid square is a mosaic rather than a homogeneous area. Therefore the assignment of each square to one of the zones is a matter of compromise, which cannot be achieved without a grain of subjectivism and uncertainty. The latter should be regarded as an inevitable cost of expert-based analyses of complex systems.

Sudnik-Wójcikowska (1987, 1988, 1992) distinguished four anthropopressure zones. We present a brief description of two of them.

> Zone A. Compact settlements prevail. Soils are ruined, rich in carbonates and nutrients; contamination of area is permanent and has existed for a long time. Ground water table is lowered and other alterations of water balance are strongly marked. Ruderal or cultivated vegetation dominates.
>
> Zone B. It is characterized by a weaker and usually shorter lasting man's impact: level of contamination is lower, disturbances in the water conditions are weaker. Suburban vegetation occurring in loose settlements, gardens and fields dominates.

Data, methods and the goal of the analysis

Sudnik-Wójcikowska (1992 and unpublished data) compiled floristic lists for 82 squares (side of ca 1 km) demarcated in eight Polish cities (Białystok, Bydgoszcz, Gdańsk, Katowice, Lublin, Lódź, Poznań, Warsaw). Each square was assigned to either zone A or B. The abundance of each species was estimated according to a 4-degree scale.

In order to eliminate rare taxa, which makes the calculations easier, the analysis was performed for species recorded in at least 10% of grid squares. The set of all taxa was, therefore, reduced to a 411-element subset denoted by C. Subsequently the whole set of objects U was split into two parts. One of them consisting of 50 objects became a training collection (set X on which the selection of attributes was performed), the rest served as a testing set, whose elements were subjected to recognition by means of fuzzy K-NN algorithm. The boolean attribute *anthropopressure* (zone A vs. zone B) was the only element of set D.

The computations have been performed using various distance measures and clustering techniques. In this paper, however, we will report only the results obtained using the Manhattan distance and the unweighted pair-group method (UPGMA). In the fuzzy K-NN algorithm, K has been set to values ranging from 1 to 15. The results for $K = 5$ will be shown below. It must be stressed, however, that very similar error rates have been obtained over a wide range of values of K.

A dendrogram computed for the entire set of 411 attributes (Fig. 6) has revealed that the squares from the same city tend to form clusters, whereas their division into two anthropopressure zones is somewhat blurred ($\gamma_C(D) = 0.63$). This observation has made it our aim to find the smallest possible collection of diagnostic species for which the quality of approximation of the partition based on *anthropopressure* would attain its maximum value.

Results and discussion

Figure 7 demonstrates the quality of approximation of R_D by R_C and the efficiency of prediction by means of fuzzy K-NN algorithm ($K = 5$) as functions of the cardinality of subsets. The values presented are averages of 50 runs. For the purpose of comparison, the results obtained for the subsets selected by GA as well as those for the subsets chosen at random are displayed in the same figure. It can be seen that in the case of GA the quality of approximation as well as the efficiency of prediction grow quickly with the increasing number of attributes and gain values close to 1 for quite small cardinalities. Table I comprises three exemplary 8-element subsets of species, denoted by C_1-C_3, for which the γ coefficient attained its maximum value of unity. These subsets also proved effective in the task of recognition of testing squares with the use of fuzzy K-NN algorithm (1 misclassified object at most in 32 subjected to recognition). In Figure 8 the dendrogram calculated for C_1 is displayed. Note that this tree contains two branches corresponding to the anthropopressure zones A and B.

Many authors have made attempts at defining some floristic parameters and their groups that could reflect the intensity of man's impact in urban agglomerations (e.g., Faliński 1971, Sukopp 1972, Kunick 1974, Kornaś 1977, Kowarik 1988, Brande et al. 1990, Pyšek & Pyšek 1990, Sudnik-Wójcikowska 1988, 1992). Although some of the indices were found to be good tools for assessing the degree of advancement of flora alteration processes, they have at least two disadvantages. First, their indicator value and usefulness in comparative studies depend strongly on definitions of species groups in the historical-geographical classification as well as particular taxon's assignment to the groups. Second, in practical tasks one cannot make use of floristic parameters unless a careful and detailed examination of the whole flora of the analyzed area is performed. This is due to the fact that both the total number of species and the number of taxa belonging to the groups are involved in the calculation of

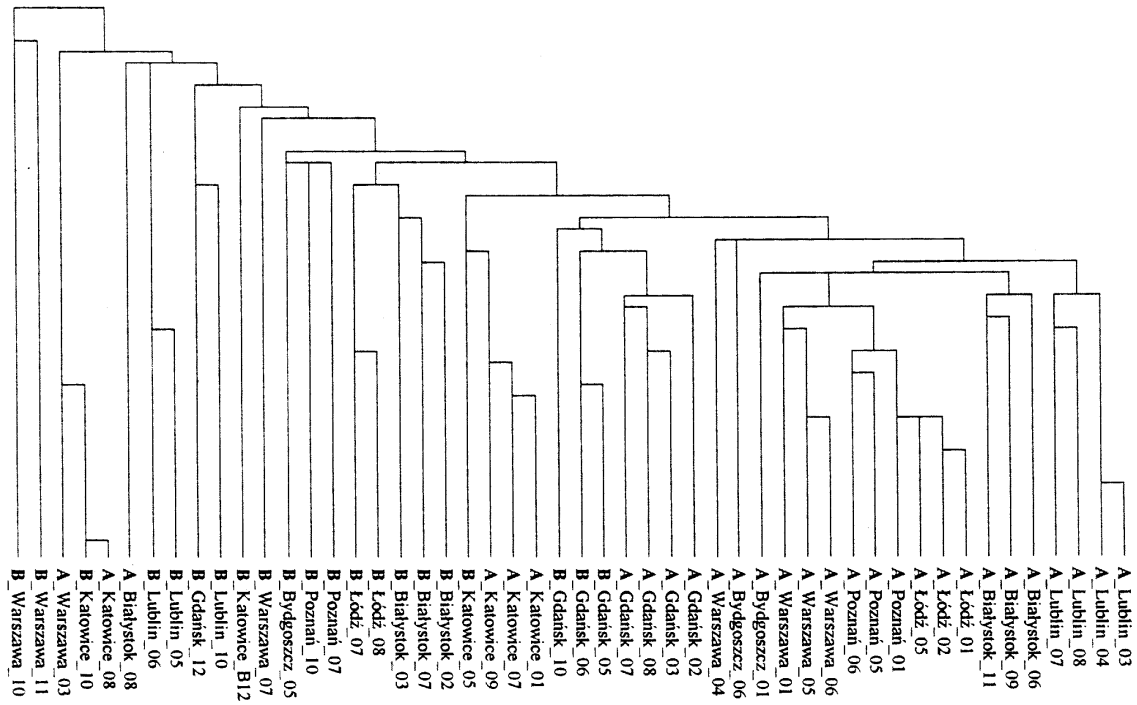


Figure 6. Dendrogram based on 411-element set *C*.

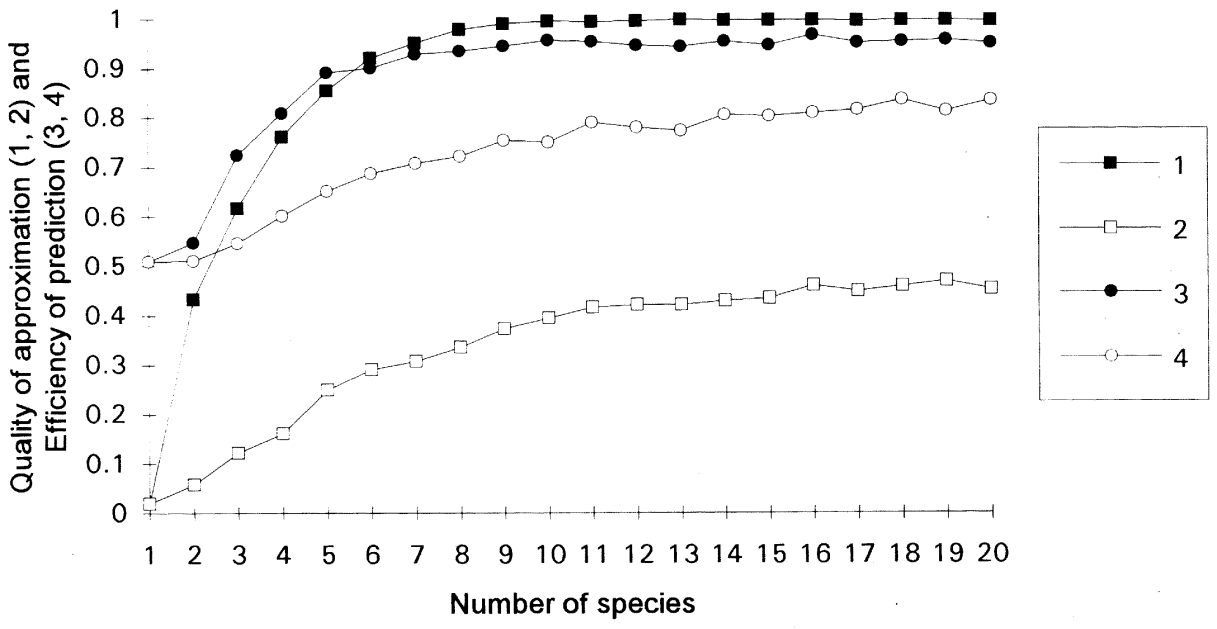


Figure 7. The quality of approximation of R_D by R_C (1, 2) and the efficiency of prediction (3, 4) vs. cardinality of subsets.
1, 3 - subsets selected with the use of a genetic algorithm,
2, 4 - subsets chosen at random.
The values presented are averages of 50 runs.

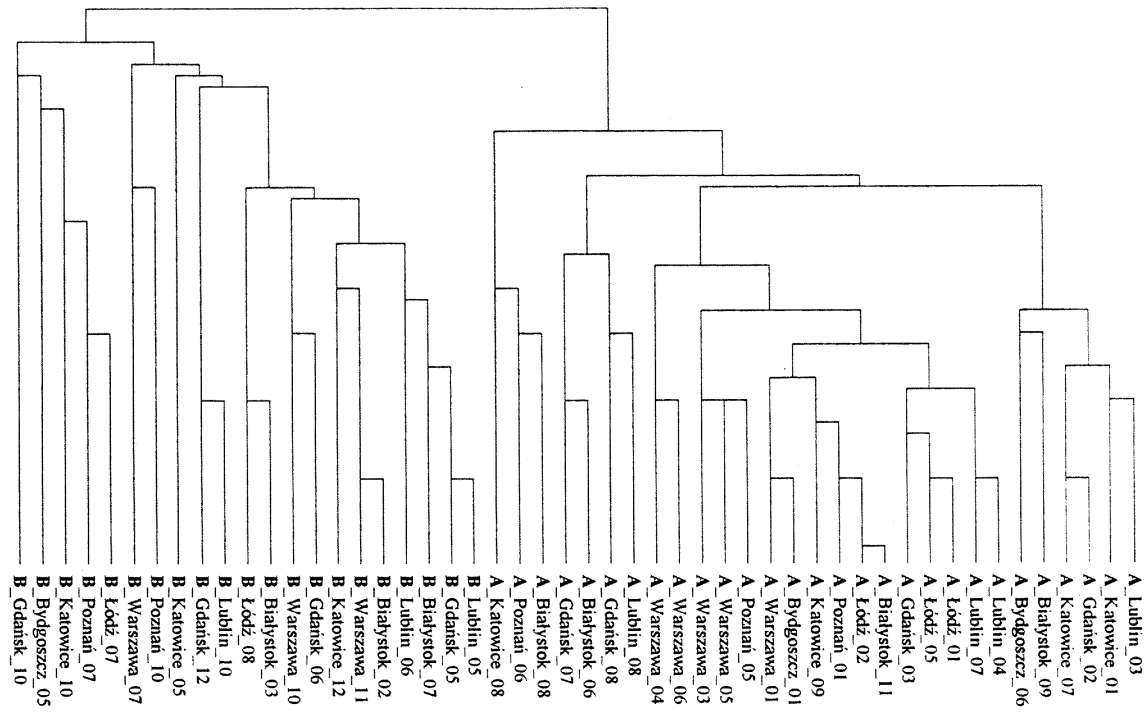


Figure 8. Dendrogram based on 8-element subset C_1.

Table 1. Three exemplary 8-element subsets of species yielding the maximum value of the quality of approximation (via Manhattan distance and UPGMA) of the crisp equivalence relation R_D (zone A vs. zone B).

	Species	Significance factor
Subset C_1	*Galium aparine*	0.43
	Epilobium angustifolium	0.38
	Spergula arvensis	0.38
	Diplotaxis muralis	0.37
	Carex hirta	0.36
	Myosotis arvensis	0.31
	Centaurea cyanus	0.25
	Centaurea scabiosa	0.22
Subset C_2	*Vicia tetrasperma*	0.33
	Apera spica-venti	0.30
	Carex hirta	0.27
	Holcus lanatus	0.25
	Pastinaca sativa	0.25
	Galium aparine	0.23
	Alliaria petiolata	0.23
	Thlaspi arvense	0.15
Subset C_3	*Vicia tetrasperma*	0.31
	Anchusa arvensis	0.29
	Tragopogon pratensis	0.26
	Galium aparine	0.25
	Hordeum distichon	0.25
	Prunella vulgaris	0.17
	Centaurea cyanus	0.16
	Raphanus raphanistrum	0.14

each parameter's value. In contrast, the approach currently proposed does not require such thorough studies. Once a rule is established it can be applied to various tasks and the only information needed is the abundance of a few species. We must admit that we make here an assumption that the decision rules obtained according to the fuzzy rough set methodology will prove to be at least so universal as the indices are. However, this is not self-evident and requires further studies.

Conclusions

An important advantage of the rough set theory as a tool for data analysis is that it is independent of any information going beyond the dataset itself. In particular, any foregoing knowledge of the statistical distributions of the attributes need not be assumed. In the fuzzified version of the theory, however, it is necessary to determine some membership functions, which in practice resolves itself into choosing an appropriate distance measure and a clustering technique. While using fuzzy rough sets has some other drawbacks (it involves computations heavier than those in the crisp case), it also has compensating advantages for it makes it possible to express a relationship between sets of attributes, interpreted as a causal nexus, a decision or diagnostic system, by means of well-known concepts (similarity, distance) and widely used methods (cluster analysis). Moreover, the theory in question furnishes a data analyst with a powerful means of reduction of data redundancy. An analysis employing the

fuzzy-rough-set-theoretical perspective can be a self-contained research procedure or serve as a starting point for advanced mathematical modeling.

Fuzzy rough sets shed new light on the problem of significance of classifications and the way in which they can be corroborated. Testing adequacy of clustering solutions involves many techniques. Among them are significance tests on external variables. Aldenderfer and Blashfield (1986) claim that "the procedures included in this category are probably among the better ways to validate a clustering solution". It seems that the degree of consistency of two partitions computed for non-overlapping subsets of attributes can be assessed in terms of quality of approximation defined by (14).

The use of a genetic algorithm in the task of finding optimal subsets makes the fuzzy rough set approach applicable to large data.

Acknowledgments: This work was supported by the State Committee for Scientific Research under a grant No. BW-1300/31/501/VII/31. I am grateful to Dr Barbara Sudnik-Wójcikowska for giving me the run of her unpublished data from eight Polish cities. I would like to express my gratitude to the Polish EMB-net node at Institute of Biochemistry and Biophysics in Warsaw for making available their computer resources. I acknowledge Prof. Andrzej Batko for his many helpful suggestions and comments on the manuscript. The assistance provided by Mrs. Izabella Krężelewska is also greatly appreciated.

References

Aldenderfer, M.S. & R.K. Blashfield. 1986. Cluster analysis. Sage Publications, London.

Allen, T.F.H. & T.B. Starr. 1982. Hierarchy: Perspectives for ecological complexity. The University of Chicago Press, Chicago.

Bauhin, C. 1622. Catalogus Plantarum circa Basileam sponte nascentium. Exemplar der Universittsbibliothek Basel , dreifach durchschlossen mit handschriftlichen Notizen von W. de Lachenal (1736-1800) 1. Aufl. Basel.

Bernhardi M., 1652. Catalogus plantarum tum exoticarum quam indigenarum quae anno MDCLI in hortis regiis Varsaviae et circa eandem in locis sylvaticis, pratensibus, arenosis et paludosis nascuntur collectarum. Gdaesk.

Bezdek, J.C. & J.D. Harris. 1978. Fuzzy partitions; an axiomatic basis for clustering. Fuzzy Sets and Systems 1: 111-127.

Bouchon-Meunier, B. 1992. Inferences with inaccuracy and uncertainty in expert systems. In: Kandel, A. (ed.). Fuzzy Expert Systems. CRC Press, Boca Raton/Ann Arbor/London/Tokyo, pp. 43-54.

Brande, A., R. Böcker, & A. Graf 1990. Changes of flora, vegetation und urban biotopes in Berlin (West). In: Sukopp, H., S. Hejny & I. Kowarik (eds.). Urban ecology. Plants and plant communities in urban environments. SPB Academic Publ., The Hague, pp. 155-165.

Brandes, D. 1989. Geographischer Vergleich der Stadtvegetation in Mitteleuropa. Braun-Blanquetia 3: 61-67.

Burton, R.M.1983. Flora of the London area. London Natural History Society, London.

Dubois, D. & H. Prade. 1990. Rough fuzzy sets and fuzzy rough sets. Int. J. of General Systems 17: 191-200.

Dubois, D. & H. Prade. 1992. Putting rough sets and fuzzy sets together. In: Slowinski, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 203-232.

Dunn, J. 1974. A graph theoretic analysis of pattern classification via Tamura's fuzzy relation. IEEE Trans. SMC: 310-313.

Faliński, J. B. 1971. Flora i rolinność synantropijna wsi i miast - próba analizy porównawczej. Mater. Zakl. Fitosoc. Stos. UW 27: 15-37.

Feoli, E. & V. Zuccarello. 1986. Ordination based on classification: yet another solution?! Abstracta Botanica 10:203-219.

Feoli, E. & V. Zuccarello. 1988. Syntaxonomy: a source of useful fuzzy sets of environmental analysis? Coenoses 3: 141-147.

Garcke, A. 1848. Flora von Halle. Halle.

Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading, MA.

Gödde, M. & R. Wittig. 1983. A preliminary attempt at a thermal division of the town of Münster (North Rhine - Westphalia, West Germany) on a floral and vegetation basis. Urban Ecol. 7: 255-262.

Grzymała-Busse, J. W. 1992. LERS - a system for learning from examples based on rough sets. In: Sowieski, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 3-18.

Hacker, G.R. 1844. Lübeckische Flora. Lübeck.

Holland, J. H. 1975. Adaptation in Natural and Artificial Systems. The University of Michigan Press, Ann Arbor.

Jackowiak, B. 1993. Atlas of distribution of vascular plants in Poznań. Prace Zakładu Taksonomii Roślin Uniwersytetu im. A. Mickiewicza w Poznaniu 2: 5-409.

Jardine, N. & R. Sibson. 1971. Mathematical Taxonomy. John Wiley & Sons Ltd., London.

Kandulski, M., J. Marciniec & K. Tukałło. 1992. Surgical wound infection - conductive factors and their mutual dependencies. In: Słowiński, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 95-110.

Keller, J.M., M.R. Gray & J.A. Givens, jr. 1985. A Fuzzy K-Nearest Neighbor Algorithm. IEEE Trans. Syst., Man, Cybern. SMC-15, 4: 580-585.

Klir, G.J. 1990. A principle of uncertainty and information invariance. Int. J. General Systems 17: 249-275.

Klett, G.T. & H.E.F. Richter. 1830. Flora der phanerogamischen Gewchse der Umgegend von Leipzig. Verl. Fr. Hofmeister, Leipzig.

Klotz, S. 1989. Merkmale der Stadtflora. Braun-Blanquetia 3: 57-60.

Klotz, S. 1990. Species/area and species/inhabitants relations in European cities. In: Sukopp H., S., Hejny & I. Kowarik (eds). Urban ecology. Plant and plant communities in urban environments. SPB Academic Publ., The Hague, pp. 99-103.

Kornaś, J. 1977. Analiza flor synantropijnych. Wiad. Bot. 21: 85-91.

Kosko, B. 1990. Fuzziness vs. probability. Int. J. General Systems 17: 211-240.

Kowarik, I. 1988. Zum menschlichen Einfluβ auf Flora und Vegetation. Theoretische Konzepte und ein Quantifizierungsansatz am Beispiel von Berlin (West). Landschaftentwicklung u. Umweltforschung 56, Tech. Univ. Berlin, Berlin.

Krysiński, J. 1992. Analysis of structure - activity relationships of quarternary ammonium compounds. In: Słowiński, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 119-136.

Kunick, W. 1974. Vernderungen von Flora und Vegetation einer Grobstadt dargestellt am Beispiel von Berlin (West). Diss. Tech. Univ. Berlin (West).

Kunick, W. 1984. Verbreitungskarten von Wildpflanzen als Bestandteil der Stadtbiotopkartierung, dargestellt am Beispiel Köln. Verh Ges. Ökologie 12: 269-275.

Landolt, E. 1991. Die Entstehung einer mitteleuropischen Stadtflora am Beispiel der Stadt Zürich. Ann. Bot. 49: 109-147.

Lucasius, C. B., A.D. Dane & G. Kateman. 1993. On *k*-medoid clustering of large data sets with the aid of a genetic algorithm: background, feasibility and comparison. Analytica Chimica Acta 282: 647-669.

McNeill, D. & P. Freiberger. 1992. Fuzzy logic. Simon & Schuster, New York/London/ Toronto /Sydney/Tokyo/Singapore.

Moraczewski, I.R. 1993a. Fuzzy logic for phytosociology 1. Syntaxa as vague concepts. Vegetatio 106: 1-11.

Moraczewski, I.R. 1993b. Fuzzy logic for phytosociology 2. Generalizations and prediction. Vegetatio 106: 13-20.

Moraczewski, I.R., B. Sudnik-Wójcikowska & W. Borkowski. Rough sets in floristic description of the inner-city of Warsaw. Flora 191, in press.

Mrózek, A. 1992. Rough sets in computer implementation of rule-based control of industrial processes. In: Słowiński, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 19-31.

Nakamura, A. 1988. Fuzzy rough sets. Note on Multiple-Valued Logic in Japan 9: 1-8.

Nakamura, A. 1992. Applications of fuzzy-rough classifications to logics. In: Słowiński, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 233-250.

Nanda, S. & S. Majumdar. 1992. Fuzzy rough sets. Fuzzy Sets and Systems 45: 157-160.

Orłowska, M.E. & M.W. Orłowski 1992. Maintenance of knowledge in dynamic information systems. In: Slowinski, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 315-329.

Pawlak, Z. 1982. Rough sets. Int. J. of Information and Computer Sciences 11: 341-356.

Pawlak, Z. 1984. Rough classification. Int. J. of Man-Machine Studies 20: 469-485.

Pawlak, Z. 1991. Rough sets. Kluwer Academic Publ., Dordrecht.

Pawlak, Z., K. Słowiński & R. Słowiński. 1986. Rough classification of patients after highly selective vagotomy for duodenal ulcer. Int. J. Man-Machine Studies 24: 413-433.

Pyšek, P. 1993. Factors affecting the diversity of flora and vegetation in central European settlements. Vegetatio 106: 89-100.

Pyšek, P. 1994. Současné metody, možnosti a omezení vyzkumu flóry a vegetace sídlišť': review. Zpr. Čes. Bot. Spolec. 29, Mater. 10: 15-32.

Pyšek, P. & A. Pyšek 1990. Comparison of the vegetation and flora of the West Bohemian villages and towns. In: Sukopp, H., S. Hejny & I. Kowarik (eds.). Urban ecology. Plants and plant communities in urban environments. SPB Academic Publ., The Hague, pp. 105-112.

Rapoport, E.H., M.E. Diaz-Betancourt & I.R.L. Lopez-Moreno. 1983. Aspectos della ecologia urbanaen la ciudad de Mexico. Pull. Limusa, Mexico, pp. 197.

Roberts, D.W. 1986. Ordination on the basis of fuzzy set theory. Vegetatio 66: 123-131.

Roberts, D.W. 1989. Fuzzy systems vegetation theory. Vegetatio 83: 71-80.

Słowiński, K., R. Słowiński & J. Stefanowski. 1988. Rough sets approach to analysis of data from peritoneal lavage in acute pancreatitis. Med. Inform. 13: 143-159.

Słowiński, R. & J. Stefanowski. 1992. 'Roughdas' and 'Roughclass' software implementations of the rough sets approach. In: Słowiński, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 445-456.

Sudnik-Wójcikowska, B. 1987. Flora miasta Warszawy i jej przemiany w ciagu XIX i XX wieku. (Flora of Warsaw and its changes during 19th and 20th centuries). Czesc 1, 2. Wyd. Uniw. Warsz., Warszawa.

Sudnik-Wójcikowska, B. 1988. Flora synanthropization and anthropopressure zones in a large urban agglomeration (exemplified by Warsaw). Flora 180: 259-265.

Sudnik-Wójcikowska, B. 1992. Studies on Indices of Flora Synanthropization. Flora 187: 37-50.

Sukopp, H. 1972. Wandel von Flora und Vegetation in Mitteleuropa under dem Einflub des Menschen. Ber. Landwirtsch. 50: 112-130.

Tadeusiewicz, R. & M. Flasiński. 1991. Pattern recognition. (In Polish). PWN, Warszawa.

Tanaka, H., H. Ishibuchi & T. Shigenaga. 1992. Fuzzy inference system based on rough sets and its application to medical diagnosis. In: Słowiński, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 111-117.

Wittig, R., D. Diesing & M. Gödde. 1985. Urbanophob - Urbanoneutral - Urbanophil. Das Verhalten der Arten gegenber dem Lebensraum Stadt. Flora 177: 265-282.

Zadeh, L.A. 1965. Fuzzy sets. Information and Control 8: 338-353.

Zadeh, L.A. 1971. Similarity relations and fuzzy orderings. Inform. Sci. 3: 177-200.

Zadeh, L.A. 1972. A new approach to system analysis. In: Marois, M. (ed.), Man and Computer, North-Holland, Amsterdam.

Zadeh, L.A. 1990. The birth and evolution of fuzzy logic. Int. J. General Systems 17: 95-105.

Ziarko, W. 1991. The Discovery, Analysis and Representation of Data Dependencies in Databases. In: Piatetsky-Shapiro, G., & W. J. Frawley (eds). Knowledge Discovery in Databases. AAAI Press/The MIT Press. Menlo Park, CA, pp. 195-209.

Ziarko, W. 1992. Acquisition of control algorithms from operation data. In: Slowinski, R. (ed.). Intelligent Decision Support, Kluwer Academic Publ., Dordrecht/Boston/London, pp. 61-75.

Manuscript received: June 1996