3

Methods of knowledge representation using rough sets

3.1 Introduction

In the physical world around us, it is impossible to find any two object (things) that are identical. By comparing any two objects, even if very similar, we will always be able to find differences between them, in particular if we consider a sufficiently large number of their features (attributes) with a sufficiently great accuracy. Of course, such a detailed description of the world is not always needed. If we decrease the precision of description, it may happen that some or even several objects that were distinguishable before become indiscernible. For example, all cities in Poland may be discernible with respect to the exact number of inhabitants. If we are interested in cities with the number of inhabitants within a given interval, e.g. from 100 to 300 thousand people, then some cities will be indiscernible with respect to the feature (attribute) "number of inhabitants". Moreover, in the description of any given object, we only consider a limited number of features, adequate to a given purpose. Quite often, we want to reduce that number to the necessary minimum. These are the problems dealt with by the theory of rough sets.

In order to facilitate further discussion, we shall introduce several notions and symbols. First, we shall define the universe of discourse U. It is the set of all objects which constitute the area of our interest. A single j-th element of this space will be denoted as x_j. Each object of the space U may be characterized using specific features. If it is a physical object, most certainly it has infinitely many features, however, we shall limit the selection to their specific subset. Let us denote the interesting set of object features of space U by the symbol Q. Let us denote the individual features by the symbol q appropriately indexed, e.g. q_i. What differentiates one object from another and makes other objects similar, these are the values of their features. Let us denote by V_q the set of values that the feature q can take. The value of feature q of the object x will be denoted as v_q^x. The vector of all object x features may be presented as $\mathbf{v}^x = \left[v_{q_1}^x, v_{q_2}^x, ..., v_{q_n}^x\right]$.

In this chapter the rough sets theory will be presented in the form of a series of definitions illustrated by examples. Table 3.1 will allow the reader an easier handling of them.

TABLE 3.1. List of definitions and examples

	Definition	Examples
3.1.	Information system	3.1, 3.2, 3.3
3.2.	Decision table	3.4
3.3.	Indiscernibility relation	
3.4.	Equivalence class	3.5, 3.6, 3.7
3.5.	Lower approximation of the set	3.8, 3.9, 3.10
3.6.	Upper approximation of the set	3.11, 3.12, 3.13
3.7.	Positive region of the set	3.8, 3.9, 3.10
3.8.	Boundary region of the set	3.14, 3.15, 3.16
3.9.	Negative region of the set	3.17, 3.18, 3.19
3.10.	$\widetilde{P}$-exactly set	3.20, 3.21, 3.22
3.11.	$\widetilde{P}$-definable set	3.20, 3.21, 3.22
3.12.	$\widetilde{P}$-accuracy of an approximation of the set	3.23, 3.24, 3.25
3.13.	Lower approximation of the family of sets	3.26
3.14.	Upper approximation of the family of sets	3.27
3.15.	Positive region of the family of sets	3.28
3.16.	Boundary region of the family of sets	3.29
3.17.	Negative region of the family of sets	3.30
3.18.	Quality of approximation of the family of sets	3.31
3.19.	Accuracy of approximation of the family of sets	3.32
3.20.	Dependence degree of attributes	3.33
3.21.	Deterministic rules	3.33
3.22.	Set of independent attributes	3.34
3.23.	Set of relatively independent attributes	3.35
3.24.	Reduct	3.36
3.25.	Relative reduct	3.36
3.26.	Indispensable attribute	3.37
3.27.	Core	3.38
3.28.	Normalized coefficient of attributes significance	3.39
3.29.	Error of approximate reduct	3.40

3.2 Basic terms

One of methods to present the information on objects characterized by the same set of features is the information system.

Definition 3.1
The information system is referred to an ordered 4-tuple $SI = \langle U, Q, V, f \rangle$ [1], where U is the set of objects, Q is the set of features (attributes), $V = \bigcup_{q \in Q} V_q$ is the set of all possible values of features, while $f : U \times Q \to V$ is called the information function. We can say that $v_q^x = f(x, q)$, of course $f(x, q) \in V_q$. The notation $v_q^x = f_x(q)$, which treats the information function as a family of functions, will be considered as equivalent. Then $f_x : Q \to V$.

Example 3.1
Let us consider a used car dealer. Currently, there are 10 cars. The universe of discourse U is therefore composed of 10 objects, which can be notated as

$$U = \{x_1, x_2, ..., x_{10}\}. \tag{3.1}$$

The car dealer notes in his documents four features of each car, which are usually referred to by customers during phone calls. These are: number of doors, horsepower, colour and make. Therefore, the set of features can be written as

$$\begin{aligned} Q &= \{q_1, q_2, q_3, q_4\} \\ &= \{\text{number of doors, horsepower, colour, make}\}. \end{aligned} \tag{3.2}$$

Based on the contents of Table 3.2 we can define the domains of particular features:

TABLE 3.2. Example of an information system

Object (U)	Number of doors (q_1)	Horsepower (q_2)	Colour (q_3)	Make (q_4)
x_1	2	60	blue	Opel
x_2	2	100	black	Nissan
x_3	2	200	black	Ferrari
x_4	2	200	red	Ferrari
x_5	2	200	red	Opel
x_6	3	100	red	Opel
x_7	3	100	red	Opel
x_8	3	200	black	Ferrari
x_9	4	100	blue	Nissan
x_{10}	4	100	blue	Nissan

$$V_{q_1} = \{2, 3, 4\}, \tag{3.3}$$

$$V_{q_2} = \{60, 100, 200\}, \tag{3.4}$$

$$V_{q_3} = \{\text{black, blue, red}\}, \tag{3.5}$$

$$V_{q_4} = \{\text{Ferrari, Nissan, Opel}\}. \tag{3.6}$$

Example 3.2
Let us consider the set of real numbers in the interval U (see Fig. 3.1), where

$$U = [0, 10). \tag{3.7}$$

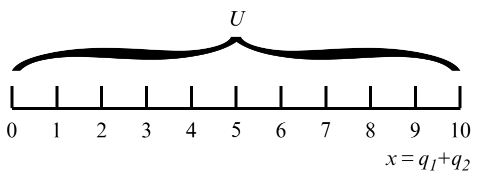

FIGURE 3.1. One-dimensional universe of discourse U

Let each element $x \in U$ be defined by two features making up the set of features

$$Q = \{q_1, q_2\}, \tag{3.8}$$

where q_1 is the integral part of the number x and q_2 is the decimal part of this number. Of course $x = \{q_1, q_2\}$. The information functions may be defined as follows

$$f_x(q_1) = \text{Ent}(x), \tag{3.9}$$

$$f_x(q_2) = x - \text{Ent}(x), \tag{3.10}$$

where function Ent$(\cdot)$ (*fr. entier*) means the integral part of the argument.

Knowing the definition of information functions, it is usually easy to define the domains of variability of particular features. In our example, they will be as follows:

$$V_{q_1} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}, \tag{3.11}$$

$$V_{q_2} = [0; 1). \tag{3.12}$$

Example 3.3
In the following example, let us consider the space of pairs

$$U = \{\mathbf{x} = [x_1; x_2] \in [0; 10) \times [0; 10)\}. \tag{3.13}$$

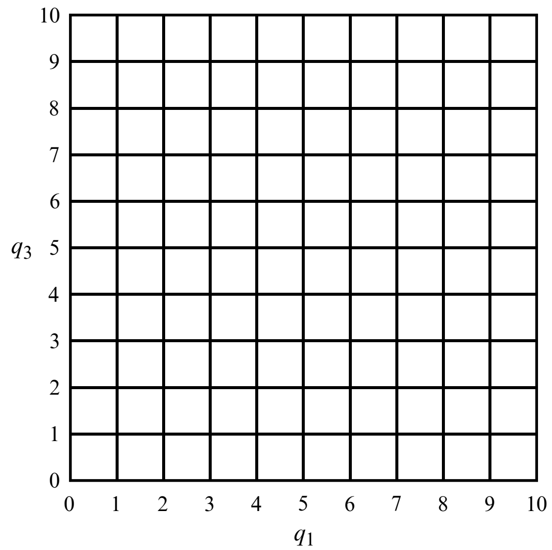

FIGURE 3.2. Two-dimensional universe of discourse U

The objects belonging to the space defined in this way may be interpreted as points located on a plane, as shown in Fig. 3.2. The most natural features of points are their coordinates x_1 and x_2. In our example, however, they will be defined otherwise. Let us define four features

$$Q = \{q_1, q_2, q_3, q_4\} \tag{3.14}$$

where q_1 is the integral part of the first coordinate of point $\mathbf{x}$, q_2 is its decimal part, and q_3 and q_4 are the integral and the decimal part of the second coordinate of the point, respectively. The information functions will therefore be defined as follows:

$$f_{\mathbf{x}}(q_1) = \mathrm{Ent}(x_1), \tag{3.15}$$

$$f_{\mathbf{x}}(q_2) = x_1 - \mathrm{Ent}(x_1), \tag{3.16}$$

$$f_{\mathbf{x}}(q_3) = \mathrm{Ent}(x_2), \tag{3.17}$$

$$f_{\mathbf{x}}(q_4) = x_2 - \mathrm{Ent}(x_2). \tag{3.18}$$

Knowing the definition of information functions, it is usually easy to define the domains of variability of particular features. In this example, they will be as follows:

$$V_{q_1} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}, \tag{3.19}$$

$$V_{q_2} = [0; 1), \tag{3.20}$$

$$V_{q_3} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}\,, \tag{3.21}$$

$$V_{q_4} = [0; 1)\,. \tag{3.22}$$

The special case of the information system is the decision table.

Definition 3.2
The decision table is the ordered 5-tuple $DT = \langle U, C, D, V, f\rangle$. The elements of the set C we call conditional features (attributes), and elements of D-decision features (attributes).

The information function f described in Definition 3.1 defines unambiguously the set of rules included in the decision table. In the notation, in the form of family of functions, the function $f_l : C \times D \to V$ defines l the decision rule of the table. The difference between the above definition and Definition 3.1 consists in separation of the set of features Q into two disjoint subsets C and D, complementary to Q. The decision tables are an alternative way of representing the information with relation to the rules:

$$R^l : \textbf{IF } c_1 = v^l_{c_1} \textbf{AND } c_2 = v^l_{c_2} \textbf{AND...AND } c_{n_c} = v^l_{c_{n_c}} \textbf{THEN } d_1 = v^l_{d_1}$$
$$\textbf{AND } d_2 = v^l_{d_2} \textbf{AND} \ldots \textbf{AND } d_{n_d} = v^l_{d_{n_d}}\,.$$

Example 3.4
Let us assume that basing on notes of the car dealer from Example 3.1, we shall build an expert system, which will define the car make based on information on the number of doors, horsepower and colour. We should divide the set Q (defined by formula (3.2)) into the set of conditional features

$$\begin{aligned} C &= \{c_1, c_2, c_3\} = \{q_1, q_2.q_3\} \\ &= \{\text{number of doors, horsepower, colour}\} \end{aligned} \tag{3.23}$$

and a single-element set of decision features

$$D = \{d_1\} = \{q_4\} = \{\text{make}\}\,. \tag{3.24}$$

Information included in the information system presented in Table 3.2 will be used to build a decision table (Table 3.3). The description of each object of space U constitutes the basis to create a single rule.

The contents included in the decision table (Table 3.3) may also be presented in the form of rules:

R^1 : **IF** $c_1 = 2$ **AND** $c_2 = 60$ **AND** $c_3 =$ blue**THEN** $d_1 =$ Nissan
R^2 : **IF** $c_1 = 2$ **AND** $c_2 = 100$ **AND** $c_3 =$ black **THEN** $d_1 =$ Nissan
...
R^{10} : **IF** $c_1 = 4$ **AND** $c_2 = 100$ **AND** $c_3 =$ blue **THEN** $d_1 =$ Nissan

TABLE 3.3. Example of the decision table

Rule (l)	Number of doors (c_1)	Horsepower (c_2)	Colour (c_3)	Make (d_1)
1	2	60	blue	Opel
2	2	100	black	Nissan
3	2	200	black	Ferrari
4	2	200	red	Ferrari
5	2	200	red	Opel
6	3	100	red	Opel
7	3	100	red	Opel
8	3	200	black	Ferrari
9	4	100	blue	Nissan
10	4	100	blue	Nissan

Now we shall present two definitions that are very important in the rough sets theory. If given two objects $x_1, x_b \in U$ have the same values of all features q belonging to the set $P \subseteq Q$, which may be notated as $\forall_q \in P$, $f_{x_a}(q) = f_{x_b}(q)$, then we say that these objects are P-indiscernible or that they are to each other in *P-indiscernibility relation* $\left(x_a, \widetilde{P} x_b\right)$.

Definition 3.3
The *P-indiscernibility relation refers* to a $\widetilde{P}$ relation defined in the space $U \times U$ satisfying

$$x_a \widetilde{P} x_b \iff \forall_q \in P;\ f_{x_a}(q) = f_{x_b}(q)\,, \tag{3.25}$$

where $x_a, x_b \in U$, $P \subseteq Q$.

It is easy to verify that the $\widetilde{P}$ relation is reflexive, symmetrical and transitive, and thus it is a relation of equivalence. The relation of equivalence divides a set in which it is defined, into a family of disjoint sets called equivalence classes of this relation.

Definition 3.4
The set of all objects $x \in U$ being in relation $\widetilde{P}$ we call *the equivalence class* of relation $\widetilde{P}$ in the space U. For each $x_a \in U$, there is exactly one such set denoted by the symbol $[x_a]_{\widetilde{P}}$, i.e.

$$[x_a]_{\widetilde{P}} = \left\{x \in U : x_a \widetilde{P} x\right\}. \tag{3.26}$$

The family of all equivalence classes of the relation $\widetilde{P}$ in the space U (called the quotient of set U by relation $\widetilde{P}$) will be denoted using the symbol P^* or $U/\widetilde{P}$.

Example 3.5

Let us define the equivalence classes of relation C-indiscernibility $\widetilde{C}$ defined by the set of features C given by formula (3.23) for the information system defined in Example 3.1:

$$[x_1]_{\widetilde{C}} = \{x_1\}\,, \tag{3.27}$$

$$[x_2]_{\widetilde{C}} = \{x_2\}\,, \tag{3.28}$$

$$[x_3]_{\widetilde{C}} = \{x_3\}\,, \tag{3.29}$$

$$[x_4]_{\widetilde{C}} = [x_5]_{\widetilde{C}} = \{x_4, x_5\}\,, \tag{3.30}$$

$$[x_6]_{\widetilde{C}} = [x_7]_{\widetilde{C}} = \{x_6, x_7\}\,, \tag{3.31}$$

$$[x_8]_{\widetilde{C}} = \{x_8\}\,, \tag{3.32}$$

$$[x_9]_{\widetilde{C}} = [x_{10}]_{\widetilde{C}} = \{x_9, x_{10}\}\,. \tag{3.33}$$

We therefore can say that the objects x_4 and x_5 are C-indiscernible, similarly to x_6 and x_7 as well as x_9 and x_{10}. The family of above specified equivalence classes will be the set

$$c^* = \{\{x_1\}\,, \{x_2\}\,, \{x_3\}\,, \{x_4, x_5\}\,, \{x_6, x_7\}\,, \{x_8\}\,, \{x_9, x_{10}\}\}\,. \tag{3.34}$$

Example 3.6

For the set of features $Q = \{q_1, q_2\}$ defined in Example 3.2 all objects are discernible, i.e. there are infinitely many one-element equivalence classes of Q-indiscernibility relation and each element of space U forms its own class. It will be different when we divide the set Q into two features sets:

$$P_1 = \{q_1\}\,, \tag{3.35}$$

$$P_2 = \{q_2\}\,. \tag{3.36}$$

For P_1-indiscernibility relation, 10 equivalence classes are formed

$$[0]_{P_1} = [0; 1)\,, \tag{3.37}$$

$$[1]_{P_1} = [1; 2)\,, \tag{3.38}$$

$$\ldots$$

$$[9]_{P_1} = [9; 10)\,. \tag{3.39}$$

Their family is the set

$$P_1^* = \{[0;1); [1;2); [2;3); [3;4); [4;5); [5;6); [6;7); [7;8); [8;9); [9;10)\}. \tag{3.40}$$

Figure 3.3 shows the exemplary equivalence class $[1]_{\widetilde{P}_1}$.

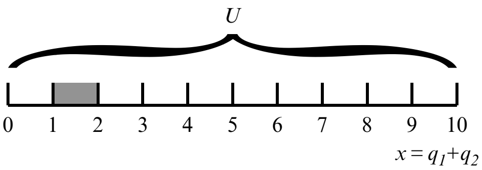

FIGURE 3.3. Example of equivalence class $[1]_{\widetilde{P}_1}$

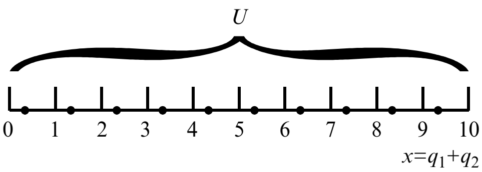

FIGURE 3.4. Example of equivalence class $[0.33]_{\widetilde{P}_2}$

For P_2-indiscernibility relation, infinitely many ten-element equivalence classes are formed. These are sets of numbers from the space U with the same decimal part

$$[x]_{P_2} = \{\widehat{x} \in U : \widehat{x} - \mathrm{Ent}\,(\widehat{x}) = x - \mathrm{Ent}\,(x)\}\,. \tag{3.41}$$

Their family is the set

$$\begin{aligned} P_2^* &= \big\{[x]_{P_2} = \{\widehat{x} \in U : \widehat{x} - \mathrm{Ent}\,(\widehat{x}) = x - \mathrm{Ent}\,(x)\} : x \in [0;1)\big\} \\ &= \big\{[x]_{P_2} = \{\widehat{x} \in U : \widehat{x} - \mathrm{Ent}\,(\widehat{x}) = x\} : x \in [0;1)\big\}\,. \end{aligned} \tag{3.42}$$

Figure 3.4 shows the exemplary equivalence class $[0.33]_{\widetilde{P}_2}$.

Example 3.7

Like in Example 3.6, for the set of features Q defined by formula (3.14) in Example 3.3, all the objects are discernible, i.e. there are infinitely many one-element equivalence classes of Q-indiscernibility relation and each element of space U forms its own class. It will be different, if we consider the features set $P \subseteq Q$ given by

$$P = \{q_1, q_3\}\,. \tag{3.43}$$

For the P-indiscernibility relation thus defined in the space U, we have 100 equivalence classes. The equivalence class of point $\mathbf{x} = (x_1, x_2)$ may be described as

$$[\mathbf{x}]_{\widetilde{P}} = \{\widehat{\mathbf{x}} = (\widehat{x}_1, \widehat{x}_2) \in U \colon \mathrm{Ent}\,(\widehat{x}_1) = \mathrm{Ent}\,(x_1) \wedge \mathrm{Ent}\,(\widehat{x}_2) = \mathrm{Ent}\,(x_2)\}\,. \tag{3.44}$$

Figure 3.5 presents the exemplary equivalence class.

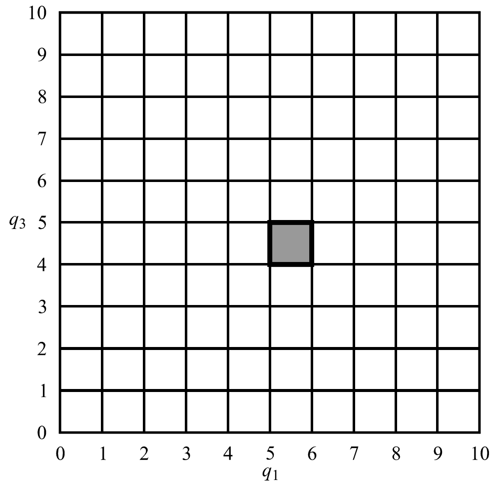

FIGURE 3.5. Example of equivalence class $[(5;4)]_{\widetilde{P}}$

Their family is the set of all square fields visible in Figs. 3.2 and 3.5. We can describe this set as follows:

$$\begin{aligned} P^* &= \left\{ [\mathbf{x}]_{\widetilde{P}} = \{\widehat{\mathbf{x}} = (\widehat{x}_1; \widehat{x}_2) \in U : \mathrm{Ent}\,(\widehat{x}_1) = \mathrm{Ent}\,(x_1) \wedge \mathrm{Ent}\,(\widehat{x}_2) \right. \\ &= \mathrm{Ent}\,(x_2) : \mathbf{x} = (x_1; x_2)\,; x_1; x_2 = 0; \ldots ; 9\} \\ &= \left\{ [\mathbf{x}]_{\widetilde{P}} = \{\widehat{\mathbf{x}} = (\widehat{x}_1; \widehat{x}_2) \in U : \mathrm{Ent}\,(\widehat{x}_1) = x_1 \wedge \mathrm{Ent}\,(\widehat{x}_2) = x_2\} : \right. \\ &\qquad \left. \mathbf{x} = (x_1; x_2)\,; x_1; x_2 = 0; \ldots ; 9 \right\}. \end{aligned} \tag{3.45}$$

3.3 Set approximation

In the space U, certain sets X may exist. We can infer that particular objects $x \in U$ belong to sets X based on the knowledge of values of their features. The set of available features $P \subseteq Q$ is usually limited and the determination of membership of the object to a specific set may not be unequivocal. This situation is described by the terms of lower and upper approximation of set $X \subseteq U$.

Definition 3.5
The set $\underline{\widetilde{P}}X$ described as follows:

$$\underline{\widetilde{P}}X = \left\{ x \in U : [x]_{\widetilde{P}} \subseteq X \right\} \tag{3.46}$$

is called $\widetilde{P}$-lower approximation of the set $X \subseteq U$.

Therefore, the lower approximation of the set X is the set of the objects $x \in U$, with relation to which on the basis of values of features P, we can certainly state that they are elements of the set X.

Example 3.8
In the space U, defined by equation (3.1) in Example 3.1 there are three sets of car makes: Ferrari, Nissan and Opel (see Table 3.2). Let us mark them with letters X_{F}, X_{N} and X_{O}:

$$X_{\mathrm{F}} = \{x_3, x_4, x_8\}\,, \tag{3.47}$$

$$X_{\mathrm{N}} = \{x_2, x_9, x_{10}\}\,, \tag{3.48}$$

$$X_{\mathrm{O}} = \{x_1, x_5, x_6, x_7\}\,. \tag{3.49}$$

We will infer the membership of various space objects based on based on the value of features of set C defined by notation (3.23). Applying directly Definition 3.5, let us determine $\widetilde{C}$-lower approximation of sets X_{F}, X_{N} and X_{O}. This definition says that the object $x \in U$ is an element of the lower approximation, if the whole equivalence class, to which it belongs, is a subset of the set X. Among the equivalence classes defined in Example 3.5, only classes $[x_3]_{\widetilde{C}}$ and $[x_8]_{\widetilde{C}}$ are the subsets of the set X_{F}, that is

$$\underline{\widetilde{C}}X_{\mathrm{F}} = \{x_3\} \cup \{x_8\} = \{x_3, x_8\}\,. \tag{3.50}$$

The object x_4 does not belong to $\underline{\widetilde{C}}X_{\mathrm{F}}$, even if it belongs to X_{F}, as the object x_5 with identical feature values from the set C, and therefore belonging to the same equivalence class, is not an element of X_{F}.

Sets $[x_2]_{\widetilde{C}}$ and $[x_9]_{\widetilde{C}} = [x_{10}]_{\widetilde{C}}$ are subsets of the set X_{N}, hence

$$\underline{\widetilde{C}}X_{\mathrm{N}} = \{x_2\} \cup \{x_9, x_{10}\} = \{x_2, x_9, x_{10}\}\,. \tag{3.51}$$

Sets $[x_1]_{\widetilde{C}}$ and $[x_6]_{\widetilde{C}} = [x_7]_{\widetilde{C}}$ are subsets of the set X_{O}, hence

$$\underline{\widetilde{C}}X_{\mathrm{O}} = \{x_1\} \cup \{x_6, x_7\} = \{x_1, x_6, x_7\}\,. \tag{3.52}$$

Example 3.9
Let us assume that in the space U defined in Example 3.2 there is a set X defined as follows:

$$X = [1,75;\ 6,50]\,. \tag{3.53}$$

Let us define the $\widetilde{P}_1$ and $\widetilde{P}_2$-lower approximation of this set. Four equivalence classes of P_1-indiscernibility relation (Example 3.6) belong entirely to the set X. Therefore, the $\widetilde{P}_1$-lower approximation will be their sum

$$\underline{\widetilde{P}_1}X = [2]_{\widetilde{P}_1} \cup [3]_{\widetilde{P}_1} \cup [4]_{\widetilde{P}_1} \cup [5]_{\widetilde{P}_1} = [2,6)\,, \tag{3.54}$$

which is illustrated in Fig. 3.6.

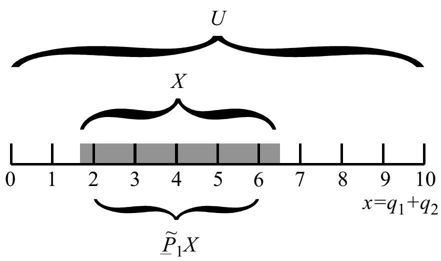

FIGURE 3.6. Lower approximation in one-dimensional universe of discourse

No equivalence class of the P_2-indiscernibility relation belongs entirely to the set X, therefore its $\widetilde{P}_2$-lower approximation is an empty set, i.e.

$$\underline{\tilde{P}_2}X = \varnothing \tag{3.55}$$

Example 3.10
Let us in the space U, defined by notation (3.13), define the set X as shown in Fig. 3.7. This figure shows the marked equivalence classes making up the $\widetilde{P}$-lower approximation of the set X. Among the 100 equivalence classes defined by formula (3.44), the lower approximation is made up by 25 equivalence classes – squares which are entirely subsets of the set X.

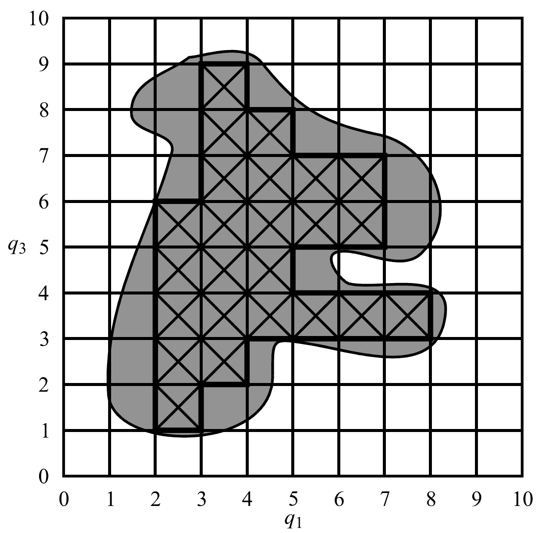

FIGURE 3.7. Lower approximation in two-dimensional universe of discourse

Definition 3.6
The set $\widetilde{P}X$ described as follows:

$$\overline{\widetilde{P}}X = \left\{x \in U : [x]_{\widetilde{P}} \cap x \neq \varnothing\right\} \tag{3.56}$$

is called $\widetilde{P}$-*upper approximation* of the set $X \subseteq U$.

The upper approximation of the set X is the set of the objects $x \in U$, with relation to which, on the basis of values of features P, we can not certainly state that they are not elements of the set X.

Example 3.11
Applying directly Definition 3.6, let us determine $\widetilde{C}$-upper approximation of sets X_{F}, X_{N}, and X_{O} defined in Example 3.8. This definition says that the object $x \in X$ is an element of the upper approximation, if the whole equivalence class, to which it belongs, has a non-empty intersection with the set X. In other words, if at least one element of a given equivalence class belongs to the set X, then each element of this equivalence class belongs to the upper approximation of the set X. Among the equivalence classes defined in Example 3.5, elements of classes $[x_3]_{\widetilde{C}}$, $[x_4]_{\widetilde{C}}$, and $[x_8]_{\widetilde{C}}$ belong to the set X_{F}, hence

$$\overline{\widetilde{C}}X_{\mathrm{F}} = \{x_3\} \cup \{x_4, x_5\} \cup \{x_8\} = \{x_3, x_4, x_5, x_8\}\,. \tag{3.57}$$

The object x_5 belongs to $\widetilde{C}$-upper approximation of the set X_{F}, even though it does not belong to X_{F}, as the object x_4 with identical values of features from set C, and therefore belonging to the same equivalence class, is an element of the set X_{F}. The set X_{N} contains elements from classes $[x_2]_{\widetilde{C}}$ and $[x_2]_{\widetilde{C}} = [x_{10}]_{\widetilde{C}}$, hence

$$\overline{\widetilde{C}}X_{\mathrm{N}} = \{x_2\} \cup \{x_9, x_{10}\} = \{x_2, x_9, x_{10}\}\,. \tag{3.58}$$

The set X_{O} contains elements from classes $[x_1]_{\widetilde{C}}$, $[x_4]_{\widetilde{C}}$ and $[x_6]_{\widetilde{C}}$, so

$$\overline{\widetilde{C}}X_{\mathrm{O}} = \{x_1\} \cup \{x_4, x_5\} \cup \{x_6, x_7\} = \{x_1, x_4, x_5, x_6, x_7\}\,. \tag{3.59}$$

Example 3.12
Let us determine $\widetilde{P}_1$ and $\widetilde{P}_2$-upper approximation of the set X defined in Example 3.9. Objects of six equivalence classes of P_1-indiscernibility relation belong to the set X. Therefore the $\widetilde{P}_1$-upper approximation will be their sum, i.e.

$$\overline{\widetilde{P}}_1X = [1]_{\widetilde{P}_1} \cup [2]_{\widetilde{P}_1} \cup [3]_{\widetilde{P}_1} \cup [4]_{\widetilde{P}_1} \cup [5]_{\widetilde{P}_1} \cup [6]_{\widetilde{P}_1} = [1, 7)\,, \tag{3.60}$$

which is illustrated by Fig. 3.8.

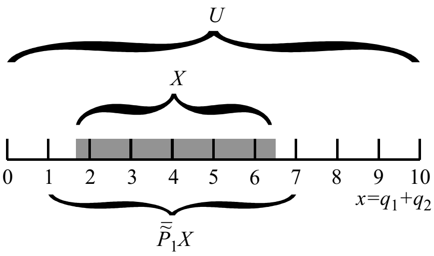

FIGURE 3.8. Upper approximation in one-dimensional universe of discourse

As the elements of all equivalence classes of the P_2-indiscernibility relation belong to the set X, so its $\widetilde{P}_2$-upper approximation is equal to the universe of discourse U, i.e.

$$\overline{\widetilde{P}}_2 X = U. \tag{3.61}$$

Example 3.13
Figure 3.9 shows the marked equivalence classes included in the P-upper approximation of the set X described in the space U defined by formula (3.13).

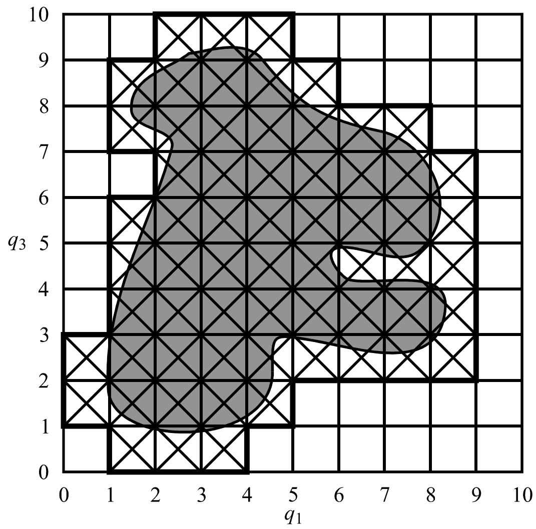

FIGURE 3.9. Upper approximation in two-dimensional universe of discourse

Definition 3.7
$\widetilde{P}$-*positive region* of the set X is defined as

$$\mathrm{Pos}_{\widetilde{P}}(X) = \underline{\widetilde{P}}X. \tag{3.62}$$

The positive region of the set X is equal to its lower approximation.

Definition 3.8
$\tilde{P}$-*boundary region* of the set X is defined as

$$\mathrm{Bn}_{\widetilde{P}}(X) = \overline{\widetilde{P}}X \setminus \underline{\widetilde{P}}X. \tag{3.63}$$

Example 3.14
By directly applying Definition 3.8, we shall find the boundary region of sets X_{F}, X_{N} and X_{O} defined in Example 3.8. We shall perform that by defining the difference of sets described in Examples 3.11 and 3.8. Therefore, we obtain

$$\begin{aligned}\mathrm{Bn}_{\widetilde{C}}(X_{\mathrm{F}}) &= \overline{\widetilde{C}}X_{\mathrm{F}} \setminus \underline{\widetilde{C}}X_{\mathrm{F}} \\ &= \{x_3, x_4, x_5, x_8\} \setminus \{x_3, x_8\} = \{x_4, x_5\},\end{aligned} \tag{3.64}$$

$$\mathrm{Bn}_{\widetilde{C}}(X_{\mathrm{N}}) = \overline{\widetilde{C}}X_{\mathrm{N}} \setminus \underline{\widetilde{C}}X_{\mathrm{N}} = \varnothing, \tag{3.65}$$

$$\mathrm{Bn}_{\widetilde{C}}(X_{\mathrm{O}}) = \overline{\widetilde{C}}X_{\mathrm{O}} \setminus \underline{\widetilde{C}}X_{\mathrm{N}} = \{x_4, x_5\}. \tag{3.66}$$

Example 3.15
Let us define the boundary region of set X defined in Example 3.9 for the set of features P_1 and P_2. In the first case, we have

$$\begin{aligned}\mathrm{Bn}_{\widetilde{P}_1}(X) &= \overline{\widetilde{P}_1}X \setminus \underline{\widetilde{P}_1}X \\ &= [1;7) \setminus [2;6) = [1;2) \cup [6;7),\end{aligned} \tag{3.67}$$

which is illustrated by Fig. 3.10. In the second case

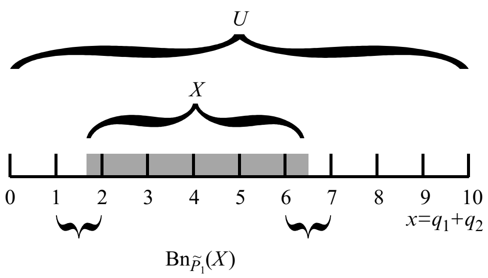

FIGURE 3.10. Boundary region in one-dimensional universe of discourse

$$\begin{aligned}\mathrm{Bn}_{\widetilde{P}_2}(X) &= \overline{\widetilde{P}_2}X \setminus \underline{\widetilde{P}_2}X \\ &= U \setminus \varnothing = U.\end{aligned} \tag{3.68}$$

Example 3.16
Figure 3.11 shows the marked equivalence classes included in the boundary region $\mathrm{Bn}_{\widetilde{P}}(X)$ described in the space U defined by formula (3.13).

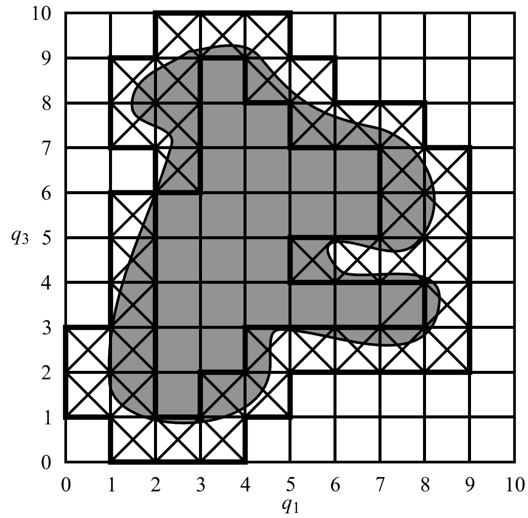

FIGURE 3.11. Boundary region in two-dimensional universe of discourse

Definition 3.9
$\widetilde{P}$-*negative region* of the set X is defined as

$$\mathrm{Neg}_{\widetilde{P}}(X) = U \setminus \overline{\overline{\widetilde{P}}}X. \tag{3.69}$$

The negative region of the set X is the set of the objects $x \in U$, with relation to which, on the basis of values of features P, we can certainly state that they are not elements of the set X.

Example 3.17
According to Definition 3.9, we shall define the negative regions of the sets X_{F}, X_{N} and X_{O} considered in Example 3.8. By defining the complement of sets defined in Example 3.11 to the space U, we shall obtain

$$\mathrm{Neg}_{\widetilde{C}}(X_{\mathrm{F}}) = U \setminus \overline{\overline{\widetilde{C}}}X = \{x_1, x_2, x_6, x_7, x_9, x_{10}\}, \tag{3.70}$$

$$\mathrm{Neg}_{\widetilde{C}}(X_{\mathrm{N}}) = U \setminus \overline{\overline{\widetilde{C}}}X = \{x_1, x_3, x_4, x_5, x_6, x_7, x_8\}, \tag{3.71}$$

$$\mathrm{Neg}_{\widetilde{C}}(X_{\mathrm{O}}) = U \setminus \overline{\overline{\widetilde{C}}}X = \{x_2, x_3, x_8, x_9, x_{10}\}. \tag{3.72}$$

Example 3.18
Let us define the boundary region of set X defined in Example 3.9 for the set of features P_1 and P_2. In the first case, we have

$$\begin{aligned}\mathrm{Neg}_{\widetilde{P}_1}(X) &= U \setminus \overline{\overline{\widetilde{P}}}_1 X \\ &= U \setminus [1;7) = [0;1) \cup [7;10),\end{aligned} \tag{3.73}$$

which is illustrated by Fig. 3.12. In the second case, we have

$$\begin{aligned}\mathrm{Neg}_{\widetilde{P}_2}(X) &= U \setminus \overline{\widetilde{P}}_2 X \\ &= U \setminus U = \varnothing.\end{aligned} \tag{3.74}$$

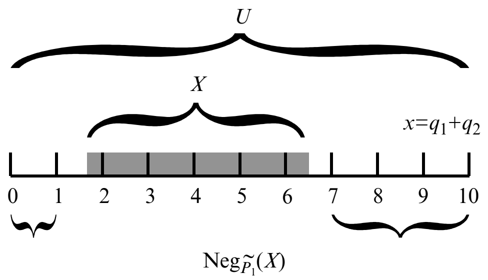

FIGURE 3.12. Negative region in one-dimensional universe of discourse

Example 3.19
Figure 3.13 shows the marked equivalence classes included in the negative region $\mathrm{Neg}_{\widetilde{P}}(X)$, determined in the space U defined by formula (3.13).

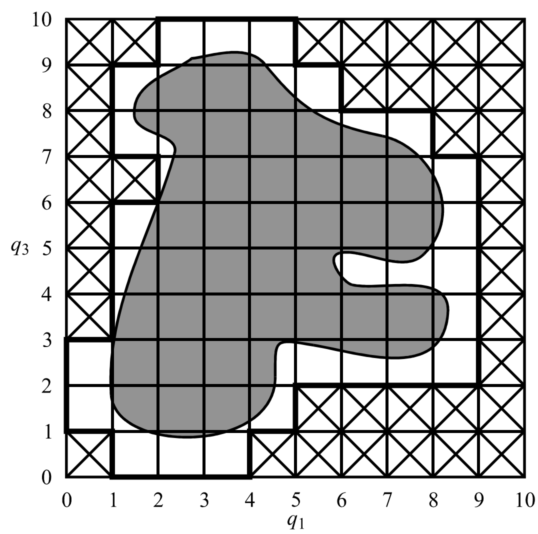

FIGURE 3.13. Negative region in two-dimensional universe of discourse

Definition 3.10
The set X is called a $\widetilde{P}$-*exactly set*, if its lower and upper approximation are equal

$$\underline{\widetilde{P}}X = \overline{\widetilde{P}}X \tag{3.75}$$

and $\widetilde{P}$-rough set otherwise

$$\underline{\widetilde{P}}X \neq \overline{\widetilde{P}}X. \tag{3.76}$$

Definition 3.11

The set X is called

a) *roughly $\widetilde{P}$-definable set*, if

$$\begin{cases} \underline{\widetilde{P}}X \neq \varnothing \\ \overline{\widetilde{P}}X \neq U, \end{cases} \tag{3.77}$$

b) *internally $\widetilde{P}$-non definable set*, if

$$\begin{cases} \underline{\widetilde{P}}X = \varnothing \\ \overline{\widetilde{P}}X \neq U, \end{cases} \tag{3.78}$$

c) *externally $\tilde{P}$-non definable set*, if

$$\begin{cases} \underline{\widetilde{P}}X \neq \varnothing \\ \overline{\widetilde{P}}X = U, \end{cases} \tag{3.79}$$

d) *totally $\widetilde{P}$-non definable set*, if

$$\begin{cases} \underline{\widetilde{P}}X = \varnothing \\ \overline{\widetilde{P}}X = U. \end{cases} \tag{3.80}$$

Example 3.20

By comparing the lower and upper approximations of sets X_{F}, X_{N} and X_{O}, described in Examples 3.8 and 3.11, we can easily notice that only the set X_{N} satisfies Definition 3.10 and is a $\widetilde{C}$-exactly set. The sets X_{F} and X_{O} satisfy equation (3.77) in Definition 3.11 and are roughly $\widetilde{C}$-definable sets, as well as $\widetilde{C}$-rough sets according to Definition 3.10.

Example 3.21

By comparing the lower and upper approximations of the set X defined in Example 3.9, determined in Examples 3.9 and 3.12, we can easily state that this set is both a $\widetilde{P}_1-$ and $\widetilde{P}_2$-rough set (Definition 3.10). Moreover, according to Definition 3.11, it is a roughly $\widetilde{P}_1$-definable set and at the same time a totally $\widetilde{P}_2$-non definable set.

Example 3.22

By analyzing Figs. 3.7 and 3.9, we can state that the set X, defined in Example 3.10 and shown in Fig. 3.7, is a $\widetilde{P}$-rough set and at the same time a $\widetilde{P}$-definable set.

Definition 3.12
The value expressed by formula

$$\mu_{\widetilde{P}}(X) = \frac{\overline{\overline{\underline{\widetilde{P}}X}}}{\overline{\overline{\overline{\widetilde{P}}X}}} \tag{3.81}$$

is called $\widetilde{P}$-*accuracy* of approximation of the set X. The symbol $\overline{\overline{A}}$ denotes the measure of the set A. In case of finite sets, we can use the cardinality as the measure, in case of continuous bounded sets, we can use such measures as the length of the interval, surface area, volume, etc.

Example 3.23
Let us determine $\widetilde{C}$-accuracy of sets X_{F}, X_{N} and X_{O}, defined in Example 3.8. By applying formula (3.81), we obtain

$$\mu_{\widetilde{C}}(X_{\mathrm{F}}) = \frac{\overline{\overline{\underline{\widetilde{C}}X_{\mathrm{F}}}}}{\overline{\overline{\overline{\widetilde{C}}X_{\mathrm{F}}}}} = \frac{2}{4} = 0.5, \tag{3.82}$$

$$\mu_{\widetilde{C}}(X_{\mathrm{N}}) = \frac{\overline{\overline{\underline{\widetilde{C}}X_{\mathrm{N}}}}}{\overline{\overline{\overline{\widetilde{C}}X_{\mathrm{N}}}}} = \frac{3}{3} = 1, \tag{3.83}$$

$$\mu_{\widetilde{C}}(X_{\mathrm{O}}) = \frac{\overline{\overline{\underline{\widetilde{C}}X_{\mathrm{O}}}}}{\overline{\overline{\overline{\widetilde{C}}X_{\mathrm{O}}}}} = \frac{3}{5} = 0.6. \tag{3.84}$$

As we can see, $\widetilde{C}$-accuracy of the approximation of the set X_{N} is 1, which confirms the previous observation that this set is a $\widetilde{C}$-exact set. In other words, it is unambiguously defined by the features belonging to the set C given by formula (3.23).

Example 3.24
In case of continuous spaces of discourses, we can define the $\widetilde{C}$-accuracy, using the length of appropriate intervals. Therefore, for the set X defined in Example 3.9, we have

$$\mu_{\widetilde{P}_1}(X) = \frac{\overline{\overline{\underline{\widetilde{P}_1}X}}}{\overline{\overline{\overline{\widetilde{P}_1}X}}} = \frac{3}{3}, \tag{3.85}$$

$$\mu_{\widetilde{P}_2}(X) = \frac{\overline{\overline{\underline{\widetilde{P}_2}X}}}{\overline{\overline{\overline{\widetilde{P}_2}X}}} = 0. \tag{3.86}$$

Example 3.25
In case of the set X, defined in Example 3.10, we can determine the $\widetilde{P}$-accuracy of approximation using the surface area, as the measure. Based on Figs. 3.7 and 3.9 we have

$$\mu_{\widetilde{P}}(X) = \frac{\overline{\overline{\underline{\widetilde{P}}X}}}{\overline{\overline{\overline{\widetilde{P}}X}}} = \frac{8}{21}. \tag{3.87}$$

3.4 Approximation of family of sets

Definitions 3.5 - 3.9 and 3.12 may be easily generalized for a certain family of sets of the space U. Let us denote the abovementioned family of sets by $\mathrm{X} = \{X_1, X_2, ..., X_n\}$.

Definition 3.13
The set $\underline{\widetilde{P}}\mathrm{X}$ described as follows:

$$\underline{\widetilde{P}}\mathrm{X} = \left\{\underline{\widetilde{P}}X_1, \widetilde{P}X_2, ..., PX_n\right\} \tag{3.88}$$

is called $\widetilde{P}$-*lower approximation* of the family of sets X.

Example 3.26
Let the elements of family of sets X be the sets X_{F}, X_{N} and X_{O}, defined in Example 3.8. We shall notate this as follows:

$$\begin{aligned} \mathrm{X} &= \{X_{\mathrm{F}}, X_{\mathrm{N}}, X_{\mathrm{O}}\} \\ &= \{\{x_3, x_4, x_8\}, \{x_2, x_9, x_{10}\}, \{x_1, x_5, x_6, x_7\}\}. \end{aligned} \tag{3.89}$$

Using the sets determined in Example 3.8, according to Definition 3.13, we can write

$$\begin{aligned} \underline{\widetilde{C}}X &= \left\{\underline{\widetilde{C}}X_{\mathrm{F}}, \underline{\widetilde{C}}X_{\mathrm{N}}, \underline{\widetilde{C}}X_{\mathrm{O}}\right\} \\ &= \{\{x_3, x_8\}, \{x_2, x_9, x_{10}\}, \{x_1, x_6, x_7\}\}. \end{aligned} \tag{3.90}$$

Definition 3.14
The set $\overline{\widetilde{P}}\mathrm{X}$ described as follows:

$$\overline{\widetilde{P}}\mathrm{X} = \left\{\overline{\widetilde{P}}\mathrm{X}_1, \overline{\widetilde{P}}\mathrm{X}_2, ..., \overline{\widetilde{P}}\mathrm{X}_n\right\} \tag{3.91}$$

is *called* $\widetilde{P}$-*upper approximation of family* of sets X.

Example 3.27
According to Definition 3.14, using the sets determined in Example 3.11, $\widetilde{C}$-upper approximation of the family of sets X defined in Example 3.26, will be

$$\begin{aligned}\overline{\widetilde{C}}\mathrm{X} &= \left\{\overline{\widetilde{C}}\mathrm{X}_{\mathrm{F}}, \overline{\widetilde{C}}\mathrm{X}_{\mathrm{N}}, \overline{\widetilde{C}}\mathrm{X}_{\mathrm{O}}\right\} \\ &= \left\{\left\{x_3, x_4, x_5, x_8\right\}, \left\{x_2, x_9, x_{10}\right\}, \left\{x_1, x_4, x_5, x_6, x_7\right\}\right\}.\end{aligned} \tag{3.92}$$

Definition 3.15
$\widetilde{P}$-*positive region of family* of the sets X is defined as

$$\mathrm{Pos}_{\widetilde{P}}\left(\mathrm{X}\right) = \bigcup_{X_i \epsilon \mathrm{X}} \mathrm{Pos}_{\widetilde{P}}\left(X_i\right). \tag{3.93}$$

Example 3.28
The $\widetilde{C}$-positive region of family of sets X, defined in Example 3.26, we can determine as follows:

$$\begin{aligned}\mathrm{Pos}_{\widetilde{C}}\left(\mathrm{X}\right) &= \mathrm{Pos}_{\widetilde{C}}\left(X_{\mathrm{F}}\right) \cup \mathrm{Pos}_{\widetilde{C}}\left(X_{\mathrm{N}}\right) \cup \mathrm{Pos}_{\widetilde{C}}\left(X_{\mathrm{O}}\right) \\ &= \left\{x_1, x_2, x_3, x_6, x_7, x_8, x_9, x_{10}\right\}.\end{aligned} \tag{3.94}$$

As it can be inferred from the example, the term of the positive region of family of sets is not equal to the term of its lower approximation – by contrast with the terms of positive region and lower approximation of sets.

Definition 3.16
$\widetilde{P}$-*boundary region of family* of the sets X is defined as

$$\mathrm{Bn}_{\widetilde{P}}\left(\mathrm{X}\right) = \bigcup_{X_i \epsilon \mathrm{X}} \mathrm{Bn}_{\widetilde{P}}\left(X_i\right). \tag{3.95}$$

Example 3.29
According to Definition 3.16, $\widetilde{C}$-boundary region of family of sets X, defined in Example 3.26, takes the form

$$\begin{aligned}\mathrm{Bn}_{\widetilde{C}}\left(\mathrm{X}\right) &= \mathrm{Bn}_{\widetilde{C}}\left(X_{\mathrm{F}}\right) \cup \mathrm{Bn}_{\widetilde{C}}\left(X_{\mathrm{N}}\right) \cup \mathrm{Bn}_{\widetilde{C}}\left(X_{\mathrm{O}}\right) \\ &= \left\{x_4, x_5\right\}.\end{aligned} \tag{3.96}$$

Definition 3.17
$\widetilde{P}$-*negative region of family* of the sets X is defined as

$$\mathrm{Neg}_{\widetilde{P}}\left(\mathrm{X}\right) = U \backslash \bigcup_{X_i \epsilon \mathrm{X}} \overline{\widetilde{P}} X_i. \tag{3.97}$$

Example 3.30
$\widetilde{C}$-negative region of family of sets X, defined in Example 3.26, according to Definition 3.17 takes the form

$$\mathrm{Neg}_{\widetilde{C}}\,(\mathrm{X}) = U \backslash \bigcup_{X_i \epsilon \mathrm{X}} \overline{\widetilde{C}} X_i = \varnothing. \tag{3.98}$$

Definition 3.18
$\widetilde{P}$-*quality of approximation of family* of sets X is determined by the expression

$$\gamma_{\widetilde{P}}\,(\mathrm{X}) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{P}}\,(\mathrm{X})}}}{\overline{\overline{U}}}. \tag{3.99}$$

Example 3.31
$\widetilde{C}$-quality of approximation of family of sets X, defined in Example 3.26, is

$$\gamma_{\widetilde{C}}\,(\mathrm{X}) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{C}}\,(\mathrm{X})}}}{\overline{\overline{U}}} = \frac{8}{10}. \tag{3.100}$$

Definition 3.19
$\widetilde{P}$-*accuracy of approximation of family* of sets X is defined by

$$\beta_{\widetilde{P}}\,(X) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{P}}\,(\mathrm{X})}}}{\sum_{X_i \epsilon \mathrm{X}} \overline{\overline{\overline{\widetilde{P}} X_i}}}. \tag{3.101}$$

Example 3.32
Using Definition 3.19 and notations (3.92) and (3.94), it is easy to check that $\widetilde{C}$-accuracy of approximation of family of sets X, defined in Example 3.26, is

$$\beta_{\widetilde{C}}\,(\mathrm{X}) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{C}}\,(\mathrm{X})}}}{\sum_{X_i \epsilon \mathrm{X}} \overline{\overline{\overline{\widetilde{C}} X_i}}} = \frac{8}{4+3+5} = \frac{2}{3}. \tag{3.102}$$

3.5 Analysis of decision tables

The theory of rough sets introduces the notion of dependency between features (attributes) of the information system. Thanks to that, we can check whether it is necessary to know the values of all features in order to unambiguously describe the object belonging to the set U.

Definition 3.20
Dependence degree of set of attributes P_2 on the set of attributes P_1, where $P_1, P_2 \subseteq Q$, is defined as follows:

$$k = \gamma_{\widetilde{P}_1}\left(P_2^*\right), \tag{3.103}$$

where $\gamma_{\widetilde{P}_1}\left(P_2^*\right)$ is determined pursuant to Definition 3.18.

The notation $P_1 \xrightarrow{k} P_2$ means that the set of attributes P_2 depends on the set of attributes P_1 to the degree $k < 1$. In case where $k = 1$, we shall simply write $P_1 \rightarrow P_2$.

The notion of dependence degree of attributes is used to define the correctness of construction of the decision table (Definition 3.2).

Definition 3.21
The rules of decision table are called *deterministic*, provided that for each pair of rules $l_a \neq l_b$ from the equality of values of all conditional attributes C, we can infer an equality of values of decision attributes D, i.e.

$$\forall_{\substack{l_a, l_b = 1, \ldots, N \\ l_a \neq l_b}} : \forall_{c \in C}\ f_{l_a}(c) = f_{l_b}(c) \rightarrow \forall_{d \in D}\ f_{l_a}(d) = f_{l_b}(d). \tag{3.104}$$

If for a certain pair of rules $l_a \neq l_b$ the above condition is not met, i.e. the equality of values of all conditional attributes C does not result in the equality of values of decision attributes D, we shall call these rules as *non-deterministic*, i.e.

$$\exists_{\substack{l_a, l_b \\ l_a \neq l_b}} : \forall_{c \in C}\ f_{l_a}(c) = f_{l_b}(c) \rightarrow \exists_{d \in D}\ f_{l_a}(d) \neq f_{l_b}(d). \tag{3.105}$$

The decision table (Definition 3.2) is *well defined*, if all its rules are deterministic. Otherwise, we say that it is *not well defined*.

Let us notice that the decision table having a set of conditional attributes C and a set of decision attributes D is well defined, if the set of decision attributes depends on the set of conditional attributes to a degree which is equal to 1 ($C \rightarrow D$), that is

$$\gamma_{\widetilde{C}}\left(D^*\right) = 1. \tag{3.106}$$

The reason for the decision table to be not well defined is that it contains the so-called non-deterministic rules. The decision table that is not well defined may be “repaired” by removing the non-deterministic rules or expanding the set of conditional attributes C.

Example 3.33
Let us examine the dependence degree of the set of attributes D on the set of attributes C defined in Example 3.4. According to Definition 3.20, we have

$$k = \gamma_{\widetilde{C}}\left(D^*\right) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{C}}\left(D^*\right)}}}{\overline{\overline{U}}}. \tag{3.107}$$

Let us notice that the equivalence classes of D-indiscernibility relation are the sets X_{F}, X_{N} and X_{O} defined in Example 3.8. Therefore

$$D^* = \{X_{\mathrm{F}}, X_{\mathrm{N}}, X_{\mathrm{O}}\}\,. \tag{3.108}$$

Based on Definition 3.15, using the lower approximations of sets X_{F}, X_{N} and X_{O} defined in Example 3.8, we obtain

$$\begin{aligned} \mathrm{Pos}_{\widetilde{C}}\,(D^*) &= \mathrm{Pos}_{\widetilde{C}}\,(X_{\mathrm{F}}) \cup \mathrm{Pos}_{\widetilde{C}}\,(X_{\mathrm{N}}) \cup \mathrm{Pos}_{\widetilde{C}}\,(X_{\mathrm{O}}) \\ &= \underline{\widetilde{C}}X_{\mathrm{F}} \cup \underline{\widetilde{C}}X_{\mathrm{N}} \cup \underline{\widetilde{C}}X_{\mathrm{O}} \\ &= \{x_3, x_8\} \cup \{x_2, x_9, x_{10}\} \cup \{x_1, x_6, x_7\} \\ &= \{x_1, x_2, x_3, x_6, x_7, x_8, x_9, x_{10}\}\,. \end{aligned} \tag{3.109}$$

By substituting the dependence (3.109) to formula (3.107), we obtain the result

$$k = \frac{\overline{\overline{\{x_1, x_2, x_3, x_6, x_7, x_8, x_9, x_{10}\}}}}{\overline{\overline{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}}}} = \frac{8}{10}. \tag{3.110}$$

We can therefore state that the set of attributes D depends on the set of attributes C to the degree $k = 0.8$, which is notated as $C \xrightarrow{0.8} D$. The obtained value $k < 1$ informs us that the decision table given in Example 3.4 is not well defined. Based on the set of conditional attributes C, we cannot unambiguously infer on the membership of objects of the space U to the particular sets X_{F}, X_{N} and X_{O}, which are equivalence classes of the D-indiscernibility relation.

The used cars dealer from Example 3.1 should expand the set of conditional attributes C, if he wants to infer unambiguously on the car make on the basis of these attributes.

The non-deterministic rules in the decision table described in Example 3.4 are the rules 4 and 5. If they are removed, a not well-defined decision table (Table 3.3) is transformed into a well defined decision table (Table 3.4).

For the decision table thus defined, it is easy to check that

$$\gamma_{\widetilde{C}}\,(D^*) = \frac{\overline{\overline{\{x_1, x_2, x_3, x_6, x_7, x_8, x_9, x_{10}\}}}}{\overline{\overline{\{x_1, x_2, x_3, x_6, x_7, x_8, x_9, x_{10}\}}}} = 1. \tag{3.111}$$

Hence, it is well defined.

TABLE 3.4. A well-defined decision table (after removing non-deterministic rules)

Rule (l)	Number of doors (c_1)	Horsepower (c_2)	Colour (c_3)	Make (d_1)
1	2	60	blue	Opel
2	2	100	black	Nissan
3	2	200	black	Ferrari
6	3	100	red	Opel
7	3	100	red	Opel
8	3	200	black	Ferrari
9	4	100	blue	Nissan
10	4	100	blue	Nissan

The second method used to create a well-defined table is to expand the sets of conditional attributes. The car dealer decided to add the type of fuel used, the type of upholstery and wheel rims, to the features considered so far. The new set of conditional attributes takes the following form

$$\begin{aligned} C &= \{c_1, c_2, c_3, c_4, c_5, c_6\} \\ &= \{\text{number of doors, horsepower, colour, fuel, upholstery, rims}\}\,. \end{aligned} \tag{3.112}$$

The domains of new features are

$$V_{C_4} = \{\text{Diesel oil, Ethyl gasoline, gas}\}\,, \tag{3.113}$$
$$V_{C_5} = \{\text{woven fabric, leather}\}\,, \tag{3.114}$$
$$V_{C_6} = \{\text{steel, aluminium}\}\,. \tag{3.115}$$

Table 3.5 presents the decision table completed with new attributes and their values.

Let us determine for decision Table 3.5 the $\widetilde{C}$-quality and the $\widetilde{C}$-accuracy of approximation of family of sets D^*. The first step is to define the family of equivalence classes of the relation $\widetilde{C}$ in the space U. Each element of the space U has at least one different value of the feature, hence

$$C^*{=}\{\{x_1\},\{x_2\},\{x_3\},\{x_4\},\{x_5\},\{x_6\},\{x_7\},\{x_8\},\{x_9\},\{x_{10}\}\}. \tag{3.116}$$

Therefore, the following positive regions of family of sets B^* are defined:

$$\mathrm{Pos}_{\widetilde{C}}\left(X_{\mathrm{F}}\right) = \underline{\widetilde{C}}X_{\mathrm{F}} = \{x_3\} \cup \{x_4\} \cup \{x_8\} = \{x_3, x_4, x_8\} = X_{\mathrm{F}}, \tag{3.117}$$

$$\mathrm{Pos}_{\widetilde{C}}\left(X_{\mathrm{N}}\right) = \underline{\widetilde{C}}X_{\mathrm{N}} = \{x_2\} \cup \{x_9\} \cup \{x_{10}\} = \{x_2, x_9, x_{10}\} = X_{\mathrm{N}}, \tag{3.118}$$

$$\begin{aligned} \mathrm{Pos}_{\widetilde{C}}\left(X_{\mathrm{O}}\right) = \underline{\widetilde{C}}X_{\mathrm{O}} &= \{x_1\} \cup \{x_5\} \cup \{x_6\} \cup \{x_7\} \\ &= \{x_1, x_5, x_6, x_7\} = X_{\mathrm{O}}. \end{aligned} \tag{3.119}$$

TABLE 3.5. A well-defined decision table (after adding attributes)

Rule (l)	Number of doors (c_1)	Horsepower (c_2)	Colour (c_3)	Fuel (c_4)	Upholstery (c_5)	Rims c_6	Make (d_1)
1	2	60	blue	Ethyl gasoline	woven fabric	steel	Opel
2	2	100	black	Diesel oil	woven fabric	steel	Nissan
3	2	200	black	Ethyl gasoline	leather	Al	Ferrari
4	2	200	red	Ethyl gasoline	leather	Al	Ferrari
5	2	200	red	Ethyl gasoline	woven fabric	steel	Opel
6	3	100	red	Diesel oil	leather	steel	Opel
7	3	100	red	gas	woven fabric	steel	Opel
8	3	200	black	Ethyl gasoline	leather	Al	Ferrari
9	4	100	blue	gas	woven fabric	steel	Nissan
10	4	100	blue	Diesel oil	woven fabric	Al	Nissan

The $\widetilde{C}$-upper approximations of these sets may be defined similarly

$$\overline{\overline{\widetilde{C}}}X_{\mathrm{F}} = \{x_3\} \cup \{x_4\} \cup \{x_8\} = \{x_3, x_4, x_8\} = X_{\mathrm{F}}, \qquad (3.120)$$

$$\overline{\overline{\widetilde{C}}}X_{\mathrm{N}} = \{x_2\} \cup \{x_9\} \cup \{x_{10}\} = \{x_2, x_9, x_{10}\} = X_{\mathrm{N}}, \qquad (3.121)$$

$$\overline{\overline{\widetilde{C}}}X_{\mathrm{O}} = \{x_1\} \cup \{x_5\} \cup \{x_6\} \cup \{x_7\} = \{x_1, x_5, x_6, x_7\} = X_{\mathrm{O}}. \qquad (3.122)$$

$\widetilde{C}$-positive region of family of sets D^* takes the form

$$\begin{aligned}\mathrm{Pos}_{\widetilde{C}}(D^*) &= \{x_3, x_4, x_8\} \cup \{x_2, x_9, x_{10}\} \cup \{x_1, x_5, x_6, x_7\} \\ &= \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8 x_9, x_{10}\} = U.\end{aligned} \qquad (3.123)$$

Now, using Definition 3.18 and 3.19, we can directly determine:

$$\gamma_{\widetilde{C}}(D^*) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{C}}(D^*)}}}{\overline{\overline{U}}} = \frac{\overline{\overline{U}}}{\overline{\overline{U}}} = \frac{10}{10} = 1, \qquad (3.124)$$

$$\beta_{\widetilde{C}}(D^*) = \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{C}}(D^*)}}}{\sum_{X_i \epsilon D^*} \overline{\overline{\overline{\overline{\widetilde{C}}}X_i}}} = \frac{10}{3+3+4} = 1. \qquad (3.125)$$

On this basis, we can unambiguously state that the decision Table 3.5 is well defined.

Definition 3.22
The set of attributes $P_1 \subseteq Q$ is *independent* in a given information system, if for each $P_2 \subset P_1$ the inequality $\widetilde{P}_1 \neq \widetilde{P}_2$ occurs. Otherwise, the set P_1 is a *dependent one*.

Example 3.34
Let us consider data given in decision Table 3.5. It is easy to notice that the set C described by formula (3.112) is a dependent set. Exemplary subsets of the set C, $C_1 = \{c_1, c_2, c_3, c_5, c_6\}$ and $C_2 = \{c_1, c_2, c_3, c_4, c_5\}$, generate such quotient of the space U, as the set C (see. 3.116). The sets C_1 and C_2 are also dependent, as the sets $C_3 = \{c_1, c_3, c_5, c_6\} \subset C_1$ and $C_4 = \{c_1, c_3, c_4, c_5\} \subset C_2$ also generate the quotient of the space U described by formula (3.116). On the other hand, the sets C_3 and C_4 are independent sets.

Definition 3.23
The set of attributes $P_1 \subseteq Q$ is *independent with respect to* the set of attributes $P_2 \subseteq Q$ (P_2-*independent*), if for each $P_3 \subset P_1$ the following inequality holds

$$\mathrm{Pos}_{\widetilde{P}_1}\left(P_2^*\right) \neq \mathrm{Pos}_{\widetilde{P}_3}\left(P_2^*\right). \tag{3.126}$$

Otherwise, the set P_1 is P_2-*dependent*.

Example 3.35
The set C_3 is an independent set in a given information system (Definition 3.22). According to Definition 3.23, it is a D-dependent set. We shall demonstrate that the set C_3 together with its subset $C_5 = \{c_1, c_3, c_6\}$ does not meet condition (3.126), i.e.

$$\mathrm{Pos}_{\widetilde{C}_3}\left(D^*\right) \neq \mathrm{Pos}_{\widetilde{C}_5}\left(D^*\right). \tag{3.127}$$

Let us notice that

$$\begin{aligned}\mathrm{Pos}_{\widetilde{C}_3}\left(D^*\right) &= \mathrm{Pos}_{\widetilde{C}_3}\left(X_{\mathrm{F}}\right) \cup \mathrm{Pos}_{\widetilde{C}_3}\left(X_{\mathrm{N}}\right) \cup \mathrm{Pos}_{\widetilde{C}_3}\left(X_{\mathrm{O}}\right) \\ &= X_{\mathrm{F}} \cup X_{\mathrm{N}} \cup X_{\mathrm{O}} = U\end{aligned} \tag{3.128}$$

and

$$\begin{aligned}\mathrm{Pos}_{\widetilde{C}_5}\left(D^*\right) &= \mathrm{Pos}_{\widetilde{C}_5}\left(X_{\mathrm{F}}\right) \cup \mathrm{Pos}_{\widetilde{C}_5}\left(X_{\mathrm{N}}\right) \cup \mathrm{Pos}_{\widetilde{C}_5}\left(X_{\mathrm{O}}\right) \\ &= X_{\mathrm{F}} \cup X_{\mathrm{N}} \cup X_{\mathrm{O}} = U\end{aligned} \tag{3.129}$$

hence

$$\mathrm{Pos}_{\widetilde{C}_3}\left(D^*\right) = \mathrm{Pos}_{\widetilde{C}_5}\left(D^*\right). \tag{3.130}$$

Definition 3.24
Every independent set $P_2 \subset P_1$ for which $\widetilde{P}_2 = \widetilde{P}_1$ is called *the reduct* of the set of attributes $P_1 \subseteq Q$.

Definition 3.25
Every P_2-independent set $P_3 \subset P_1$ for which $\widetilde{P}_3 = \widetilde{P}_1$ is called *the relative reduct* of a set of attributes $P_1 \subseteq Q$ with respect to P_2 (the so-called P_2-reduct).

Example 3.36
If we return to the discussions in Examples 3.34 and 3.35; we can notice that the sets C_3 and C_4 presented therein are the reducts of the set C, whereas the set C_5 is the D-reduct of the set C_3 and C.

Definition 3.26
The attribute $p \in P_1$ is *indispensable from* P_1, if for $P_2 = P_1 \backslash \{p\}$, the equation $\widetilde{P}_2 \neq \widetilde{P}_1$ holds. Otherwise, the attribute p is *dispensable.*

Example 3.37
Using Definition 3.26, we shall check the indispensability of particular attributes $c \in C$ in the information system forming the decision Table 3.5. It is easy to check that

$$\begin{aligned} C^* = \{&\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}\}, \end{aligned} \tag{3.131}$$

$$\begin{aligned} (C\backslash \{c_1\})^* = \{&\{x_1\}, \{x_2\}, \{x_3, x_8\}, \{x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_9\}, \{x_{10}\}\} \neq C^*, \end{aligned} \tag{3.132}$$

$$\begin{aligned} (C\backslash \{c_2\})^* = \{&\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}\} = C^*, \end{aligned} \tag{3.133}$$

$$\begin{aligned} (C\backslash \{c_3\})^* = \{&\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}\} \neq C^*, \end{aligned} \tag{3.134}$$

$$\begin{aligned} (C\backslash \{c_4\})^* = \{&\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}\} = C^*, \end{aligned} \tag{3.135}$$

$$\begin{aligned} (C\backslash \{c_5\})^* = \{&\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}\} = C^*, \end{aligned} \tag{3.136}$$

$$\begin{aligned} (C\backslash \{c_6\})^* = \{&\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \\ &\{x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}\}\} = C^*. \end{aligned} \tag{3.137}$$

As we can see, the attributes c_1 and c_3 are indispensable, while the attributes c_2, c_4, c_5 and c_6 are superfluous.

Definition 3.27
The set of all indispensable attributes from the set P is called a *core* of P, which is notated as follows:

$$\mathrm{CORE}\,(P) = \left\{ p \in P : \widetilde{P}' \neq \widetilde{P}, P' = P \backslash \{p\} \right\}. \tag{3.138}$$

Example 3.38
Using the results from Example 3.37, we can define the core of the set of attributes C as

$$\mathrm{CORE}\,(C) = \{c_1, c_3\}\,. \tag{3.139}$$

Definition 3.28
The normalized coefficient of significance of subset of the set of conditional attributes $C' \subset C$ is expressed by the following formula

$$\sigma_{(C,D)}\,(C') = \frac{\gamma_{\widetilde{C}}\,(D^*) - \gamma_{\widetilde{C''}}\,(D^*)}{\gamma_{\widetilde{C}}\,(D^*)}, \tag{3.140}$$

where $C'' = C \setminus C'$. Of course, in a special case the set C' may be a one-element set, then the coefficient (3.140) will express the significance of one conditional attribute.

The coefficient of significance plays an important role in the analysis of decision tables. The zero value obtained for a given subset of conditional attributes C indicates that this subset may be deleted from the set of conditional attributes without any detriment to the approximation of family of sets D^*.

Example 3.39
Let us determine the significance of an exemplary subset of the set of conditional attributes C defined by the notation (3.112). In Example 3.33, we have demonstrated (formula (3.124)), that $\widetilde{C}$-quality of approximation of family of sets D^* for a well-defined decision table amounts to 1. For $C' = \{c_1\}$, we have $C'' = \{c_2, c_3, c_4, c_5, c_6\}$ and

$$\begin{aligned} \gamma_{\widetilde{C''}}\,(D^*) &= \frac{\overline{\overline{\mathrm{Pos}_{\widetilde{C''}}\,(D^*)}}}{\overline{\overline{U}}} \\ &= \frac{\overline{\overline{X_{\mathrm{F}} \cup X_{\mathrm{N}} \cup X_{\mathrm{O}}}}}{\overline{\overline{U}}} = 1. \end{aligned} \tag{3.141}$$

Hence

$$\sigma_{(C,D)}\,(\{c_1\}) = \frac{1-1}{1} = 0. \tag{3.142}$$

Therefore, the attribute c_1 in the given decision table is insignificant, and due to that, its removal will not impact the quality of approximation of family of sets D^*.

For $C' = \{c_4, c_5, c_6\}$, we get $C'' = \{c_1, c_2, c_3\}$, hence

$$\gamma_{\widetilde{C''}}(D^*) = \frac{\overline{\overline{\{x_3, x_8\} \cup \{x_2, x_9, x_{10}\} \cup \{x_1, x_6, x_7\}}}}{\overline{\overline{U}}} = \frac{8}{10}, \tag{3.143}$$

which, after substituting to formula (3.140), gives the value

$$\sigma_{(C,D)}(\{c_4, c_5, c_6\}) = \frac{1 - 0.8}{1} = 0.2. \tag{3.144}$$

Based on the above discussion we see the attributes c_4, c_5 and c_6 added in Example 3.33 (Table 3.5) are of low significance.

Definition 3.29
Any given subset of the set of conditional attributes $C' \subset C$ is called a rough D-reduct of the set of attributes C, and *the approximation error* of this reduct is defined as follows:

$$\varepsilon_{(C,D)}(C') = \frac{\gamma_{\widetilde{C}}(D^*) - \gamma_{\widetilde{C'}}(D^*)}{\gamma_{\widetilde{C}}(D^*)}. \tag{3.145}$$

Example 3.40
Let us determine an approximation error of the set $C' = \{c_1, c_2, c_3\}$ which is the rough D-reduct of set of attributes C (decision Table 3.5). Using the result (3.143), we have

$$\varepsilon_{(C,D)}(\{c_1, c_2, c_3\}) = \frac{1 - 0.8}{1} = 0.2. \tag{3.146}$$

3.6 Application of LERS software

LERS (*Learning from Examples based on Rough Sets*) software [67] has been created by RS Systems company. Its task is to generate the rule base, based on examples entered and to test the rule base generated or prepared independently. The data entered may be subject to some initial processing, among others by removing contradictions, eliminating or completing missing data and quantization of numerical values.

In order to present the capabilities of LERS software, let us consider two cases of data analysis. The first case, already discussed in the Example 3.1 – it is the case of the used car dealer. The second case – the problem of classification of Iris flowers, an example often used to illustrate and compare the performance of computational intelligence algorithms.

Example 3.41 (Cars in the parking lot)
Let the decision table describing the used car dealer from Example 3.1 have the form as in Table 3.5. In order to have a clear presentation of this example, it has been presented again in Table 3.6.

TABLE 3.6. Original decision table (before reduction)

Number of doors (c_1)	Horsepower (c_2)	Colour (c_3)	Fuel (c_4)	Upholstery (c_5)	Rims (c_6)	Make (d_1)
2	60	blue	Ethyl gasoline	woven fabric	steel	Opel
2	100	black	Diesel oil	woven fabric	steel	Nissan
2	200	black	Ethyl gasoline	leather	Al	Ferrari
2	200	red	Ethyl gasoline	leather	Al	Ferrari
2	200	red	Ethyl gasoline	woven fabric	steel	Opel
3	100	red	Diesel oil	leather	steel	Opel
3	100	red	gas	woven fabric	steel	Opel
3	200	black	Ethyl gasoline	leather	Al	Ferrari
4	100	blue	gas	woven fabric	steel	Nissan
4	100	blue	Diesel oil	woven fabric	Al	Nissan

In order to enter the data from Table 3.6 to LERS software, the following file must be prepared:

$<$ a, a, a a a a d $>$

[doors	horsepower]	colour	fuel	upholstery	rims	make
2	60	blue	Ethyl gasoline	woven fabric	steel	Opel
2	100	black	Diesel oil	woven fabric	steel	Nissan
2	200	black	Ethyl gasoline	leather	alum	Ferrari
2	200	red	Ethyl gasoline	leather	alum	Ferrari
2	200	red	Ethyl gasoline	woven fabric	steel	Opel
3	100	red	Diesel oil	leather	steel	Opel
3	100	red	gas	woven fabric	steel	Opel
3	200	black	Ethyl gasoline	leather	alum	Ferrari
4	100	blue	gas	woven fabric	steel	Nissan
4	100	blue	Diesel oil	woven fabric	alum	Nissan

In the first row of the file, the division to conditional attributes (a) and decision attributes (d) has been made. The second row contains the names of particular attributes. Based on data entered, LERS software generated 5 rules containing 8 conditions in total:

IF rims is steel **AND** colour is red **THEN** make is Opel
IF horsepower is 60 **THEN** make is Opel
IF doors is 4 **THEN** make is Nissan
IF colour is black **AND** fuel is Diesel oil **THEN** make is Nissan
IF horsepower is 200 **AND** upholstery is leather **THEN** make is Ferrari

TABLE 3.7. Decision table after removing redundant data

Number of doors (c_1)	Horsepower (c_2)	Colour (c_3)	Fuel (c_4)	Upholstery (c_5)	Rims (c_6)	Make (d_1)
	60					Opel
		black	Diesel oil			Nissan
	200			leather		Ferrari
	200			leather		Ferrari
		red			steel	Opel
		red			steel	Opel
		red			steel	Opel
	200			leather		Ferrari
4						Nissan
4						Nissan

The process of rules generation may be interpreted as removing redundant data from the decision table, which is shown in Table 3.7. The algorithm used for rules generation and removal of redundant data uses the rough sets theory.

By removing repeating entries from 3.7, we obtain Table 3.8, identical with the generated set of rules.

TABLE 3.8. Decision table obtained after reduction

Number of doors (c_1)	Horsepower (c_2)	Colour (c_3)	Fuel (c_4)	Upholstery (c_5)	Rims	Make (d_4)
		red			steel	Opel
	60					Opel
4						Nissan
		black	Diesel oil			Nissan
	200			leather		Ferrari

Example 3.42 (Classification of Iris flowers)
As it has been mentioned before, the problem of classification of Iris flowers is often used as an example to illustrate the performance of different types of computational intelligence algorithms. The task consists in determining the membership of flowers to one of three classes: Setosa, Virginica and Versicolor. The decision is made based on the value of four conditional attributes describing the dimensions (length and width) of the leaf and the flower petal.

We have 150 samples in our disposal, including 147 unique ones (not recurrent); 50 of them belongs to each of three classes. Table 3.9 presents the ranges of variability of particular attributes.

TABLE 3.9. Ranges of variability of attributes (classification of Iris flowers)

Attribute	Range	Number of unique values
p1	$\langle 4.3; 7.9 \rangle$	35
p2	$\langle 2.0; 4.4 \rangle$	23
p3	$\langle 1.0; 6.9 \rangle$	43
p4	$\langle 0.1; 2.5 \rangle$	22
Iris	Setosa, Virginica, Versicolor	3

The data have been divided into a learning and a testing part; 40 samples from each class have been selected randomly for the learning part and the remaining 30 samples have been used to create the testing part. Based on the contents of the learning sequence, the input file for LERS software has been prepared in the form:

```
< a   a   a   a   d >
[p1   p2   p3  p4 iris]
 4.4  2.9  1.4  0.2  Setosa
 4.8  3.0  1.4  0.1  Setosa
 5.4  3.4  1.7  0.2  Setosa
 ...
```

Based on data entered, LERS generated 34 rules containing 41 conditions in total:

IF p4 is 0.2 THEN iris is Setosa
IF p4 is 0.4 THEN iris is Setosa
IF p4 is 0.3 THEN iris is Setosa
IF p4 is 0.1 THEN iris is Setosa
IF p4 is 0.5 THEN iris is Setosa
IF p4 is 0.6 THEN iris is Setosa
IF p4 is 1.3 THEN iris is Versicolor
IF p4 is 1.5 AND p3 is 4.5 THEN iris is Versicolor

IF p4 is 1.0 THEN iris is Versicolor
IF p4 is 1.4 THEN iris is Versicolor
IF p4 is 1.5 AND p3 is 4.9 THEN iris is Versicolor
IF p4 is 1.2 THEN iris is Versicolor
IF p4 is 1.5 AND p1 is 5.9 THEN iris is Versicolor
IF p3 is 4.7 THEN iris is Versicolor
IF p4 is 1.1 THEN iris is Versicolor
IF p3 is 4.8 AND p1 is 5.9 THEN iris is Versicolor
IF p3 is 4.6 THEN iris is Versicolor
IF p4 is 1.6 AND p1 is 6.0 THEN iris is Versicolor
IF p4 is 2.1 THEN iris is Virginica
IF p4 is 2.3 THEN iris is Virginica
IF p3 is 5.5 THEN iris is Virginica
IF p4 is 2.0 THEN iris is Virginica
IF p1 is 7.3 THEN iris is Virginica
IF p3 is 6.0 THEN iris is Virginica
IF p3 is 5.1 THEN iris is Virginica
IF p3 is 5.8 THEN iris is Virginica
IF p3 is 6.1 THEN iris is Virginica
IF p4 is 2.4 THEN iris is Virginica
IF p4 is 1.8 AND p1 is 6.2 THEN iris is Virginica
IF p4 is 1.7 THEN iris is Virginica
IF p3 is 6.7 THEN iris is Virginica
IF p3 is 5.0 THEN iris is Virginica
IF p3 is 5.7 THEN iris is Virginica
IF p3 is 4.9 AND p2 is 2.7 THEN iris is Virginica

In the next step, data included in the learning sequence have been quantized so that the corresponding decision table remained still deterministic. LERS software defined the intervals given in Table 3.10 for particular conditional attributes.

The original input file has been replaced with the file presented below. Each value of the decision attribute has been replaced with an interval identifier it belongs to.

```
! Decision table produced by LERS (C version 1.0)
! First the attribute names list ...
!
[p1  p2  p3  p4  iris]
!
! Now comes the actual data. Please note that one example
! does NOT necessarily occupy one physical line
!
4.4..5.05  2.75..2.95  1..2.6  0.1..0.8  Setosa
4.4..5.05  2.95..3.05  1..2.6  0.1..0.8  Setosa
5.05..5.65  3.25..3.45  1..2.6  0.1..0.8  Setosa
...
```

TABLE 3.10. Result of quantization (classification of iris flowers)

Attribute	Range	Number of samples
p1	$\langle 4.4; 5.05 \rangle$	26
	$\langle 5.05; 5.65 \rangle$	26
	$\langle 5.65; 6.15 \rangle$	22
	$\langle 6.15; 6.65 \rangle$	22
	$\langle 6.65; 7.9 \rangle$	24
p2	$\langle 2; 2.75 \rangle$	24
	$\langle 2.75; 2.95 \rangle$	18
	$\langle 2.95; 3.05 \rangle$	21
	$\langle 3.05; 3.25 \rangle$	21
	$\langle 3.25; 3.45 \rangle$	17
	$\langle 3.45; 4.4 \rangle$	19
p3	$\langle 1; 2.6 \rangle$	40
	$\langle 2.6; 4.85 \rangle$	40
	$\langle 4.85; 6.9 \rangle$	40
p4	$\langle 0.1; 0.8 \rangle$	40
	$\langle 0.8; 1.65 \rangle$	42
	$\langle 1.65; 2.5 \rangle$	38
Iris	Setosa	40
	Virginica	40
	Versicolor	40

Based on the file so prepared, LERS software generated 11 rules containing altogether 41 conditions:

IF p3 is > **THEN** iris is Setosa
IF p4 is <0.8; 1.65> **AND** p3 is <2.6; 4.85> **THEN** iris is Versicolor
IF p2 is <3.05; 3.25> **AND** p1 is <5.65; 6.15> **THEN** iris is Versicolor
IF p4 is <0.8; 1.65> **AND** p2 is <3.05; 3.25> **THEN** iris is Versicolor
IF p1 is <6.15; 6.65> **AND** p2 is <2; 2.75> **AND** p4 is <0.8; 1.65> **THEN** iris is Versicolor
IF p3 is <4.85; 6.9> **AND** p4 is <1.65; 2.5> **THEN** iris is Virginica
IF p3 is <4.85; 6.9> **AND** p2 is <2.75; 2.95> **THEN** iris is Virginica
IF p1 is <5.65; 6.15> **AND** p3 is <4.85; 6.9> **THEN** iris is Virginica
IF p4 is <1.65; 2.5> **AND** p2 is <2.75; 2.95> **THEN** iris is Virginica
IF p2 is <2.95; 3.05> **AND** p1 is <6.65; 7.9> **THEN** iris is Virginica
IF p2 is <2; 2.75> **AND** p4 is <1.65; 2.5> **THEN** iris is Virginica

The rules obtained in the first and in the second trial have been used to classify the samples included in the testing set. The results of both experiments have been presented in Table 3.11. The first four columns contain the values of conditional attributes for test samples, the fifth column contains the correct result (decision attribute), the sixth column is the result obtained using the first set of rules, and the seventh column is the result obtained using the second set of rules.

By analyzing Table 3.11, one can notice, among others things, that the initial quantization of data, which resulted in the set of rules operating on intervals, leads to a more efficient inference system.

TABLE 3.11. Results of classification of iris flowers

p1	p2	p3	p4	pattern	classification 1	classification 2
5.0	3.6	1.4	0.2	Setosa	Setosa	Setosa
4.9	3.1	1.5	0.1	Setosa	Setosa	Setosa
4.3	3.0	1.1	0.1	Setosa	Setosa	Setosa
5.0	3.0	1.6	0.2	Setosa	Setosa	Setosa
5.5	4.2	1.4	0.2	Setosa	Setosa	Setosa
5.1	3.4	1.5	0.2	Setosa	Setosa	Setosa
5.1	3.8	1.5	0.3	Setosa	Setosa	Setosa
5.1	3.5	1.4	0.3	Setosa	Setosa	Setosa
4.6	3.1	1.5	0.2	Setosa	Setosa	Setosa
5.1	3.8	1.9	0.4	Setosa	Setosa	Setosa
5.1	2.5	3.0	1.1	Versicolor	Versicolor	Versicolor
6.1	2.8	4.7	1.2	Versicolor	Versicolor	Versicolor
6.0	2.7	5.1	1.6	Versicolor	???	Virginica
5.5	2.4	3.8	1.1	Versicolor	Versicolor	Versicolor
4.9	2.4	3.3	1.0	Versicolor	Versicolor	Versicolor
6.7	3.0	5.0	1.7	Versicolor	Virginica	Virginica
6.2	2.2	4.5	1.5	Versicolor	Versicolor	Versicolor
6.8	2.8	4.8	1.4	Versicolor	Versicolor	Versicolor
5.7	2.8	4.5	1.3	Versicolor	Versicolor	Versicolor
5.8	2.6	4.0	1.2	Versicolor	Versicolor	Versicolor
6.3	2.5	5.0	1.9	Virginica	Virginica	Virginica
6.1	3.0	4.9	1.8	Virginica	???	Virginica
6.3	2.9	5.6	1.8	Virginica	???	Virginica
6.7	3.1	5.6	2.4	Virginica	Virginica	Virginica
5.8	2.8	5.1	2.4	Virginica	Virginica	Virginica
6.1	2.6	5.6	1.4	Virginica	Versicolor	Virginica
6.4	2.7	5.3	1.9	Virginica	???	Virginica
6.9	3.1	5.4	2.1	Virginica	Virginica	Virginica
6.0	3.0	4.8	1.8	Virginica	???	???
6.4	2.8	5.6	2.2	Virginica	???	Virginica

3.7 Notes

The theory of rough sets was created by professor Zdzisław Pawlak [161 - 164]. The definitions provided in this chapter, as well as various applications of rough sets, are presented in a monograph [140], which is the first more comprehensive study on this subject in the Polish language. We refer the Reader interested in various aspects of rough sets to a rich set of publications [66, 67, 158, 177, 180, 233]. In Section 3.6, the LERS software has been used to generate the rules using the rough sets method. This software has been kindly made available for the purposes of this publication by professor Jerzy Grzymała-Busse of Kansas University, USA.