

Ivan Kramosil

An alternative approach to rough sets

*Kybernetika*, Vol. 28 (1992), No. 1, (1)--25

Persistent URL: <http://dml.cz/dmlcz/124969>

## Terms of use:

© Institute of Information Theory and Automation AS CR, 1992

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these

*Terms of use.*



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library*  
<http://project.dml.cz>

PE 4582/28. 1992. + přel.

AN ALTERNATIVE APPROACH TO ROUGH SETS

IVAN KRAMOSIL



01848/01

From the classical point of view, as conceived by Z. Pawlak, rough sets are generated by classical subsets of a basic space and by an indiscernibility relation defined on this space; this relation expresses limited abilities to discern the elements which are members of a given classical set from the other ones. In this paper an alternative approach is suggested and investigated, when the possibility to decide about the validity of a membership predicate for a given element of the basic space is qualitatively and quantitatively classified and only the subsets for which this degree of decidability is below a given threshold value are taken into consideration. Some special quantification relations as well as relations between rough sets under the alternative approach, Dempster-Shafer theory and fuzzy sets are briefly discussed.

1. FROM INDISCERNIBILITY RELATIONS TO ROUGH SETS

As far as their historical origins are concerned, the notion of rough set should be considered as the secondary one, following the primary and gnoseologically motivated notion of indiscernibility relation. Take a non-empty basic set (space, universe of discourse)  $X$ ; from the classical set-theoretic point of view elements of  $X$  are equipped with the full individual identity, are strictly separable and, in fact, separated from each other. In other words said, if  $x, y \in X$  and  $x = y$ , then  $x$  and  $y$  are even theoretically inseparable and  $x, y$  are nothing else than two names for one and the same entity. On the other handside, if  $x, y \in A$  and  $x \neq y$ , then  $x$  and  $y$  have nothing in common, in Leibniz terms, the monade of each  $x \in X$  is either this  $x$  itself, or the singleton set  $\{x\}$ , the difference being only a matter of technical convenience (at least from the point of view of our purposes in this paper when we have no needs to go into the details of type theory).

Evidently, from a more practical viewpoint such a trivial and idealized solution to the discernibility problem for elements of a given set is very far from being the satisfactory one. As a first approximation we may admit that there is an equivalence relation  $\approx$  defined on  $X$ , which is called and interpreted as an *indiscernibility relation* on  $X$ . Hence, if  $x, y \in X$  and  $x \approx y$ , we are not able, within the supposed and limited scope of our abilities, to discern  $x$  from  $y$ , i. e., to decide, whether  $x = y$  in the classical sense of identity relation or whether  $x \neq y$ .

Let us recall the definition of equivalence relation according to which, for all  $x, y, z \in X$ , (1)  $x \approx x$ , (2) if  $x \approx y$ , then  $y \approx x$ , and (3) if  $x \approx y$  and  $y \approx z$ , then  $x \approx z$ . The conditions

rapšat do  
naa / mesery

(1) and (2) seem to be obvious for each reasonable indiscernibility relation, but the transitivity condition (3) can be taken as a rather strong idealization and subjected to a discussion. Namely, (3) excludes the possibility that there is a finite sequence  $x_1, x_2, \dots, x_n$  of elements of  $X$  such that  $x_1$  is discernible from  $x_n$ , but each  $x_i$  is indiscernible from  $x_{i+1}$ . So, the well-known “monkey-Darwin” paradox still remains paradoxical under this formalization of indiscernibility relation (just to recall the matter: ancestors – parents – of each human being are human beings, descendants of each monkey are monkeys, there are finite generations from a monkey to Darwin, and still Darwin is, as a man, discernible from each monkey). However, more detailed and sophisticated reasonings about these problems would bring us into the area of alternative set theory and this ultimately is not our aim. So, we shall accept, in what follows, the reduction of indiscernibility relations to appropriate equivalence relations.

For a fixed indiscernibility (equivalence) relation  $\approx$  on  $A$  denote by  $[x]$  the equivalence class in  $X/\approx$  containing a given  $x \in X$ . Now, for each  $V \subset X$  two sets can be defined, namely,

$$\begin{aligned} V_* &= \bigcup \{[x] : [x] \subset V\} = \{x \in V : (\forall y \in X)(y \approx x \Rightarrow y \in V)\}, & (1) \\ V^* &= \bigcup \{[x] : [x] \cap V \neq \emptyset\} = \{x \in X : (\exists y \in V)(x \approx y)\}. & (2) \end{aligned}$$

Evidently,  $V_* \subset V \subset V^* \subset X$ . The pair  $\langle V_*, V^* \rangle$  is called the *rough set* generated in  $X$  by  $V$  and  $\approx$ , cf., e.g., the papers of Kubát [4], Orłowska [9] – [11] and Pawlak [12] – [14] in the references below for more details on rough sets defined through indiscernibility relations. The notion of rough set can be easily extended to each pair  $\langle V_*, V^* \rangle$ ,  $V_* \subset V^* \subset X$ , of sets, in other words, each such a pair of sets can be taken as the rough set generated by a set  $V$  such that  $V_* \subset V \subset V^*$  and by the equivalence relation  $\approx_0$  such that  $X/\approx_0 = \{V_*, V^* - V_*, X - V^*\}$ , i.e.,  $x \approx_0 y$  iff either  $x, y \in V_*$ , or  $x, y \in V^* - V_*$ , or, finally,  $x, y \in X - V^*$ .

The intuition and interpretation behind these notions is as follows. Given  $V \subset X$  and  $x \in X$ , we are not able to decide directly, whether  $x \in V$  or not, the only we can do is to take the class  $[x]$  and to decide, whether  $[x] \subset V$  or whether  $[x] \cap V \neq \emptyset$ . So,  $V_*$  is the set of all elements which can be claimed to be in  $V$  without any doubts, as all  $[x]$  is in  $V$ , and  $X - V^*$  is the set of all elements which can be surely claimed not to be in  $V$ . After all, for elements from  $V^* - V_*$  we are in doubts, their membership in  $V$  cannot be decided within the given framework of our limited abilities. The same interpretation can be applied to each pair  $\langle V_*, V^* \rangle$  of sets such that  $V_* \subset V \subset V^* \subset X$ , even not mentioning explicitly the indiscernibility relation and speaking only about our limited abilities to decide the membership relation for  $V$ . This interpretation will be discussed in more details in the next chapter.

In general, rough sets can be understood as a set-theoretic analogy of the well-known three-valued Łukasiewicz logic with the third truth value interpreted as “I do not know...”, “we are not able to decide...” or, less subjectively, “it is not known...”

which of the two classical truth values is the correct one for the proposition in question. Let us also emphasize the fact that in this approach, indiscernibility relations and rough sets are conceived as purely *gnoseological* notions. Hence, our *abilities* to discern between elements and to decide membership relations are taken as *limited* and *relativized*, but no doubts are admitted as far as the existence of *all* subsets of the space  $X$  and the ideal separability and disjointness of any two elements of  $X$  are considered. Admitting such doubts, we shall arrive, in what follows, to an alternative and *ontologically* based approach to the notion of rough sets.

## 2. ONTOLOGIZATION OF ROUGH SETS

Aiming to outline a qualitatively different and more ontological in its nature approach to the notion of rough sets, we have to begin with the very axiomatics of the set theory. We do not repeat here this axiomatics (in its Zermelo–Fraenkel setting, to be more correct) in all details and refer to, e. g., Kuratowski–Mostowski’s monography [6] for mathematical details and to Maddy’s essay [7] for interesting philosophical discussions, also Frankel and Bar-Hillel [2] monograph may be of interest in this context (cf. the list of references below). In the center of our attention will lie two axioms, namely *axiom of power set* (PSA) and *axiom of extensionality* (EA). Just to recall both of them, PSA claims:

“If  $A$  is a set, then the collection  $\mathcal{P}(A)$  of all subsets of  $A$  is also a set.” And, for EA we have:

“If  $A$  is a set and  $P$  is a unary predicate, i. e., a well-formed formula of a formalized language  $\mathcal{L}$  with single free indeterminate, then the collection  $A|P$  of all elements of  $A$  for which  $P$  holds, is also a set.”

As a rule, we write

$$A|P = \{x : x \in A, P(x)\}. \quad (3)$$

Let us remember that, not taking into consideration the axiom of choice, just PSA and EA are, because of their non-constructive features, probably the most often discussed, objected, and criticized axioms of ZF-set theory. Or, it is just PSA which is, in the greatest degree, “responsible” for the “explosion” of cardinals in ZF-set theory. On the other hand, an unrestricted definition of new subsets through unary predicates has given arise to the paradoxa like “the set of all sets” (applying the predicate “to be a set” to the universum of all entities).

What is important in our context is the purely platonist and ideal character of existence of sets created due to both the axioms above. The sets  $\mathcal{P}(A)$  and  $A|P$ , defined by PSA and EA, are ascribed *the same* kind or modus of existence like the original set  $A$ , and are supposed to be able to be subjected to the same operations as  $A$  itself, i. e., we may obtain the new sets  $\mathcal{P}(\mathcal{P}(A))$  and  $(A|P)|Q$ , if  $Q$  is a unary predicate. *The possible intractability of these sets* (because of a great cardinality of  $\mathcal{P}(A)$ , say, or because of a very subtle nature of the predicate  $P$ ) *does not influence, anyhow, their status of*

*existence.* E. g., the set of Goedel numbers of the true sentences of Peano arithmetic “exists” in the same sense as each, even the most simple one, recursive subset of the set  $\mathcal{N}$  of all non-negative integers.

Due to a continual repetition in all textbooks, monographs, courses of elementary set theory and mathematics, etc., all what we have said till now seems to be very trivial, self-evident, and not worth repeating. On the other hand side, many interesting scientific discoveries originate from “putting into question” the seemingly most self-evident assumptions of the current scientific paradigma [5]. Trying to do so in the case of EA we shall arrive at rough sets in another way than that described above.

As before, let  $X$  be a non-empty basic set (space, universe of discourse), let  $B$  be another non-empty set, let  $\preceq$  be a *partial-ordering relation* defined on  $B$ . I. e., for each  $x, y, z \in B$ ,  $x \preceq x$ ,  $x \preceq y$  and  $y \preceq x$  imply  $x = y$  in the sense of the identity relation on  $B$  and, finally,  $x \preceq y$  and  $y \preceq z$  imply  $x \preceq z$ . Let  $\mathcal{L}$  be a formalized first-order language with the set  $\underline{\text{Pred}}$  of unary predicates (well-formed formulas with a single free indeterminate). Finally, let

$$\rho : \mathcal{P}(X) \times \underline{\text{Pred}} \times X \rightarrow B \quad (4)$$

be a *partial mapping* (called *reference system*) ascribing to (some, in general) subsets  $A \subset X$ , predicates  $P$ , and elements  $x \in X$  the value  $\rho(A, P, x)$  from  $B$ .

**Definition 1.** Given  $A \subset X$ ,  $P \in \underline{\text{Pred}}$ , and  $b \in B$ , the (classical) set  $A|P$  is called a *b-subset of the set A*, if for all  $x \in A$ ,  $\rho(A, P, x)$  is defined and

$$\rho(A, P, x) \preceq b. \quad (5)$$

$B$  is a *b-subset of A*, if there is  $P \in \underline{\text{Pred}}$  such that  $B$  is  $A|P$  and  $A|P$  is a *b-subset of A*.

The intuition and interpretation behind is as follows:  $\rho(A, P, x)$  express the degree of effort and expenses necessary to decide, *within a given scope of abilities corresponding to the reference system  $\rho$* , whether  $P(x)$  hold supposing we know that  $x$  is in  $A$ . When  $\rho(A, P, x)$  is not defined, it is beyond the abilities of the reference system in question to decide, whether  $P(x)$  holds or not, even when knowing that  $x \in A$ . The degrees of effort and expenses are supposed to be at least partially comparable and a *b-subset of a set* is such one for which the corresponding membership relation defined by a predicate can be decided for all elements within uniformly majorized limits of efforts (complexity, cost,...).

In general, there is no reason to expect that any classical subset of the space  $X$  will be also its *b-subset* for a given  $(B, \preceq)$ ,  $b$  and  $\rho$  (cf. the next chapters for more details). So, we may approximate classical subsets  $V$  of  $X$  by “upper” and “lower” *b-subsets*, i. e. by pairs  $(V_*, V^*)$  of *b-subsets of X* such that  $V_* \subset V \subset V^*$ . We may also either ascribe the status of existence only to *b-subsets of X*, taking a fixed  $b \in B$ , or less radically, to introduce different modes of existence for *b-subsets of X* (the “stronger” or intuitionistic

existence), and for classical subsets of  $X$  (the “weaker” or platonistic existence). We can, again, focus our attention to pairs  $(V_*, V^*)$  of “strongly existing” subsets, i. e. to  $b$ -subsets of  $X$  such that  $V_* \subset V^*$ . So we have arrived back to the rough sets, moreover, this interpretation can be easily ontologized, as the existence of  $b$ -subsets and rough sets does not depend on and does not descend from classical sets and indiscernibility relations on them, hence, rough sets can be taken as ontologically primary entities.

### 3. MONOTONOUS REFERENCES SYSTEMS

Let us define, in the partially ordered set  $\langle B, \preceq \rangle$ , two binary operations  $\vee$  (supremum) and  $\wedge$  (infimum) by the following conditions valid for all  $x, y, z \in B$ :

$$x \wedge y \preceq x \preceq x \vee y, \quad x \wedge y \preceq y \preceq x \vee y, \quad (6)$$

$$\begin{aligned} x \preceq z \text{ and } y \preceq z & \text{ implies } x \vee y \preceq z, \\ z \preceq x \text{ and } z \preceq y & \text{ implies } z \preceq x \wedge y. \end{aligned} \quad (7)$$

Of course, both the operations  $\vee$  and  $\wedge$  are partial, but supposing that  $x \wedge y$  and  $x \vee y$  are defined, they are defined uniquely. Or, if some  $q \in B$  satisfies the conditions (6) and (7) for  $x \wedge y$ , then  $x \wedge y \preceq q$ , but also  $q \preceq x \wedge y$ , so that  $q = x \wedge y$ , and similarly for  $x \vee y$ . In order to keep the  $\wedge$  and  $\vee$  symbols unambiguously free for this use, let us denote by  $\underline{\text{et}}$  and  $\underline{\text{vel}}$  the functions of conjunction and disjunction of the language  $\mathcal{L}$ , so that if  $P, Q$  are unary predicates from  $\underline{\text{Pred}}$  with the same free indeterminate, then  $P \underline{\text{et}} Q$  and  $P \underline{\text{vel}} Q$  are also unary predicates from  $\underline{\text{Pred}}$ .

**Definition 2.** The reference system  $\rho$  is called *monotonous* on  $A \subset X$ , if for each  $P, Q \in \underline{\text{Pred}}$  and  $x \in A$  such that the expressions below are defined,

$$\rho(A, P \underline{\text{et}} Q, x) = \rho(A, P, x) \vee \rho(A, Q, x), \quad (8)$$

$$\rho(A, P \underline{\text{vel}} Q, x) = \rho(A, P, x) \wedge \rho(A, Q, x). \quad (9)$$

The system  $\rho$  is called *weakly monotounous*, if for  $P, Q, x$  as above

$$\rho(A, P \underline{\text{vel}} Q, x) \preceq \rho(A, P, x) \preceq \rho(A, P \underline{\text{et}} Q, x), \quad (10)$$

the same relation for  $Q$  immediately follows when interchanging  $P$  and  $Q$  in (10).

**Fact 1.** (a) If the reference system  $\rho$  is weakly monotounous, then for all  $b \in B$  the system  $\mathcal{F}(b, A, \rho)$  of all  $b$ -subsets of  $A$  is closed with respect to finite unions.  
 (b) If the reference system  $\rho$  is monotounous, then  $\mathcal{F}(b, A, \rho)$  is closed with respect to finite unions and joints.

**Proof.** (a) Let  $B, C \in \mathcal{F}(b, A, \rho)$ , then there exist  $P, Q \in \underline{\text{Pred}}$  such that  $B = A|P, C = A|Q$ , and  $\rho(A, P, x) \preceq b, \rho(A, Q, x) \preceq b$  for all  $x \in A$ . Then, due to (10),  $\rho(A, P \underline{\text{vel}} Q, x) \preceq b$  as well, and  $B \cup C = A|(P \underline{\text{vel}} Q)$ , so that  $B \cup C \in \mathcal{F}(b, A, \rho)$ .

(b) If  $\rho$  is monotonous, then it is, clearly, also weakly monotonous. Moreover,  $\rho(A, P \underline{\text{et}} Q, x) = \rho(A, P, x) \vee \rho(A, Q, x)$  is majorized by  $b$  due to the definition of  $\vee$ , cf. (9), but  $B \cap C = A \mid (P \underline{\text{et}} Q)$ , so that  $B \cap C \in \mathcal{F}(b, A, \rho)$ .  $\square$

Having decided the validity of  $P(x)$  for some  $P \in \underline{\text{Pred}}$  and  $x \in A$ , we have at hand also the decision about the validity of its negation  $\underline{\text{non}} P(x)$ , so that the assumption

$$\rho(A, P, x) = \rho(A, \underline{\text{non}} P, x) \quad (11)$$

seems to be quite natural. However, combining (11) with the definition of monotonous or ever weakly monotonous reference systems, we arrive at some consequences which look rather unnatural. E. g., (8), (9) and (11) yield that

$$\rho(A, P \underline{\text{vel}} \underline{\text{non}} P, x) = \rho(A, P \underline{\text{et}} \underline{\text{non}} P, x) = \rho(A, P, x), \quad (12)$$

and (10) combined with (11) yield that

$$\rho(A, P, x) \preceq \rho(A, P \underline{\text{et}} \underline{\text{non}} P, x). \quad (13)$$

Both (12) and (13) contradicts the intuition that it should be very simple to decide about the validity of a tautology, or about the non-validity of a contradiction, for each  $x \in A$  and that the complexity of this decision should not depend on particular  $P \in \underline{\text{Pred}}$ .

Let us recall that  $\mathcal{L}$  is the formalized language (the first-order one) the set of unary predicates of which is  $\underline{\text{Pred}}$ . Let  $\mathcal{T} \subset \mathcal{L}$  be a theory defined in  $\mathcal{L}$  in such a way that  $\mathcal{T}$  is the set of all logical consequences of a recursive subset  $Ax \subset \mathcal{L}$ , the formulas from  $Ax$  are called *axioms*. As a rule, we write  $\vdash x$  instead of  $x \in \mathcal{T}$ . If  $\Rightarrow$  and  $\Leftrightarrow$  are the implication and equivalence functors of the language  $\mathcal{L}$ , we may define a binary relation  $\equiv$  on the set  $\underline{\text{Sent}}$  of all sentences of  $\mathcal{L}$  (i. e. closed formulas without free indeterminates) as follows: for each  $X, Y \in \underline{\text{Sent}}$ ,

$$X \equiv Y \quad \text{iff} \quad \vdash_{\text{ar}} X \Leftrightarrow Y. \quad (14)$$

Evidently,  $\equiv$  is an equivalence relation on  $\mathcal{L}$ , so that we may generate the quotient algebra  $\underline{\text{Sent}} / \equiv$ . It is nothing else than the well-known *Lindenbaum-Tarski algebra* generated by the language  $\mathcal{L}$  and theory  $\mathcal{T}$ , therefore, we shall write  $LT(\mathcal{L}, \mathcal{T})$  instead of  $\underline{\text{Sent}} / \equiv$ .

Set, for  $X, Y \in \underline{\text{Sent}}$ ,

$$X \preceq_* Y \quad \text{iff} \quad \vdash_{\text{ar}} X \Rightarrow Y. \quad (15)$$

Evidently,  $\preceq_*$  is a partial ordering on  $\underline{\text{Sent}}$ , which is invariant with respect to the  $\equiv$  relation (i. e., if  $X \preceq_* Y$ ,  $X \equiv X_1$ , and  $Y \equiv Y_1$ , then  $X_1 \preceq_* Y_1$ ). Hence,  $\preceq_*$  uniquely defines the partial ordering relation  $\preceq_{LT}$  on  $LT(\mathcal{L}, \mathcal{T})$ . For  $X \in \underline{\text{Sent}}$  we denote by  $|X|$  the class in  $LT(\mathcal{L}, \mathcal{T})$  which contains  $X$ . Let  $P \in \underline{\text{Pred}}$ , let  $x \in A$ , let  $\mathcal{L}$  contain an individual constant  $\bar{x}$  the fixed interpretation of which is the element  $x$ . E. g., if  $A$  is the set  $\mathcal{N}$  of all non-negative integers,  $n \in \mathcal{N}$ , then  $\bar{n}$  is the *numeral* corresponding to  $n$  (cf. e. g., Kleene). Now,  $P(\bar{x})$  is a sentence of  $\mathcal{L}$ , so that  $|P(\bar{x})| \in LT(\mathcal{L}, \mathcal{T})$ .

Suppose that there is  $\bar{x}$  in  $\mathcal{L}$  for each  $x \in A$  and that  $\mathcal{L}$  contains the membership predicate  $\in$ . Consider the reference system  $\rho_0$  defined on  $\mathcal{P}(X) \times \underline{\text{Pred}} \times X$  and taking its values in the partially ordered set  $\langle LT(\mathcal{L}, T), \preceq_{LT} \rangle$ ;  $\rho_0$  is defined by

$$\rho_0(A, P, x) = |\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x})|. \quad (16)$$

**Fact 2.** The reference system  $\rho_0$  is monotonous. For each  $A \subset X$  and  $P \in \underline{\text{Pred}}$ ,  $A|P$  is a  $|\underline{\text{non}}(\forall y \in A)P(y)|$ -subset of  $A$  (with respect to  $\rho_0$ ).

**Proof.** Using some well-known first-order predicate calculus tautologies we obtain that

$$\begin{aligned} \rho_0(A, P \text{ et } Q, x) &= & (17) \\ &= |\bar{x} \in A \text{ et } (\underline{\text{non}}(P \text{ et } Q))(\bar{x})| = \\ &= |\bar{x} \in A \text{ et } ((\underline{\text{non}} P) \text{vel} (\underline{\text{non}} Q))(\bar{x})| = \\ &= |\bar{x} \in A \text{ et } ((\underline{\text{non}} P)(\bar{x}) \text{vel} (\underline{\text{non}} Q)(\bar{x}))| = \\ &= |\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x}) \text{ vel } (\bar{x} \in A \text{ et } \underline{\text{non}} Q(\bar{x}))| = \\ &= |\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x})| \vee_{LT} |\bar{x} \in A \text{ et } \underline{\text{non}} Q(\bar{x})| = \\ &= \rho_0(A, P, x) \vee_{LT} \rho_0(A, Q, x), \end{aligned}$$

where  $\vee_{LT}$  is the supremum operation in  $LT(\mathcal{L}, T)$  defined by  $\preceq_{LT}$ . Similarly,

$$\begin{aligned} \rho_0(A, P \text{ vel } Q, x) &= & (18) \\ &= |\bar{x} \in A \text{ et } (\underline{\text{non}}(P \text{ vel } Q))(\bar{x})| = \\ &= |\bar{x} \in A \text{ et } \underline{\text{non}}(P(\bar{x}) \text{ vel } Q(\bar{x}))| = \\ &= |\bar{x} \in A \text{ et } (\underline{\text{non}} P(\bar{x}) \text{ et } \underline{\text{non}} Q(\bar{x}))| = \\ &= |\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x}) \text{ et } (\bar{x} \in A \text{ et } \underline{\text{non}} Q(\bar{x}))| = \\ &= |\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x})| \wedge_{LT} |\bar{x} \in A \text{ et } \underline{\text{non}} Q(\bar{x})| = \\ &= \rho_0(A, P, x) \wedge_{LT} \rho_0(A, Q, x), \end{aligned}$$

where  $\wedge_{LT}$  is the infimum operation in  $LT(\mathcal{L}, T)$  defined by  $\preceq_{LT}$ . Hence, the reference system  $\rho_0$  is monotonous.

Let  $P \in \underline{\text{Pred}}$ ,  $A \subset X$ ,  $x \in X$ , then a well-known first-order predicate tautology yields that

$$\vdash ((\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x})) \Rightarrow \underline{\text{non}}(\forall y) (y \in A \Rightarrow P(y))), \quad (19)$$

or, using the common conventions concerning the bounded (or restricted) quantifiers,

$$\vdash ((\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x})) \Rightarrow \underline{\text{non}}(\forall y \in A)P(y)). \quad (20)$$

Hence,

$$|\bar{x} \in A \text{ et } \underline{\text{non}} P(\bar{x})| \preceq_{LT} |\underline{\text{non}}(\forall y \in A)P(y)|, \quad (21)$$

so that, due to (16)

$$\rho_0(A, P, x) \preceq_{LT} |\underline{\text{non}}(\forall y \in A)P(y)| \quad (22)$$

for each  $x \in A$ . So,  $A|P$  is a  $|\underline{\text{non}}(\forall y \in A)P(y)|$ -subset of  $A$  with respect to  $\rho_0$ .  $\square$

## 4. NUMERICAL REFERENCE SYSTEMS

When considering the most intuitive and natural particular cases of reference systems, those with numerical values of the function  $\rho$  would deserve at least the same degree of attention as the boolean-valued ones mentioned in Chapter 3. As the most simple case let us consider the well-ordered set  $(\mathcal{N}, \leq)$  where  $\mathcal{N} = \{0, 1, 2, \dots\}$  is the set of non-negative integers (in their standard interpretation, to avoid any misunderstandings) and  $\leq$  is the usual ordering relation on  $\mathcal{N}$ . In this chapter we shall investigate reference systems  $\rho$  such that  $\rho(A, P, x) \in \mathcal{N}$ ; the value  $\rho(A, P, x)$  can be interpreted as, say, the number of steps which must be executed by a testing oracle in order to decide whether  $P(x)$  holds or not, supposing that  $x \in A$ . E. g., if  $P$  is a recursive predicate relative to  $A$ , i. e.,  $A|P$  is a recursive subset of  $A$ , then  $\rho(A, P, x)$  may be defined as the number of steps taken by an appropriate Turing machine computing the truth value of  $P(x)$ .

Generalizing these informal ideas and giving them a more formalized covering we arrive at the following model.

Let  $\mathcal{A}$  be a finite set (alphabet) consisting of abstract elements (letters); the well-known examples of alphabets are the binary, decadic or English (26 letters) ones. The set of  $\mathcal{A}^*$  of all finite sequences of letters, including the empty sequence  $\Lambda$ , will play the role of the basic space  $X$  from above, hence,  $X = \mathcal{A}^* = \bigcup_{n=0}^{\infty} \mathcal{A}^n$ ,  $\mathcal{A}^0 = \{\Lambda\}$ . Let  $A$  be a fixed subset of  $\mathcal{A}^*$ , let  $\mathcal{L}_1$  be a language the individual indeterminates of which are supposed to range over  $\mathcal{A}^*$ , hence, if  $\underline{\text{Pred}}_1$  is the set of unary predicates of  $\mathcal{L}_1$  and  $P \in \underline{\text{Pred}}_1$ , then for each  $x \in \mathcal{A}^*$ ,  $P(x)$  is either true or false. Let  $\mathcal{U} = \langle U_1, U_2, \dots \rangle$  be a finite or infinite sequence of subsets of  $\mathcal{A}^*$ , for our sakes  $\mathcal{U}$  can be always taken as infinite possibly with  $U_i = \mathcal{A}^*$  for all  $i$  large enough. *Turing machine with the oracle  $\mathcal{U}$*  ( $\mathcal{U}$ -TM, abbreviately) over the alphabet  $\mathcal{A}$  is a Turing machine which is able to execute, besides the usual actions, also this one:

Having reached (one of) certain specific instantaneous description(s)  $\alpha q_i \beta$  (cf. Davis [1], e. g., for definition), the machine produces, first of all, a word  $x = x(\alpha q_i \beta) \in \mathcal{A}^*$ , uniquely and effectively defined by the instantaneous description  $\alpha q_i \beta$ . Then, the machine divides this word into the pair  $\langle x_1, x_2 \rangle \in \mathcal{A}^* \times \mathcal{A}^*$ , using a fixed decomposition function. Finally, the machine computes the index  $i(x_1)$  of  $x_1$  with respect to a fixed ordering of  $\mathcal{A}^*$ , say, the well-known lexicographical one, and decides, whether  $x_2 \in U_{i(x_1)}$  or not. If  $x_2 \in U_{i(x_1)}$ , the machine enters the given internal state  $q_j$ , if it is not the case, it enters  $q_k \neq q_j$ , in both the cases continuing its work according to the definition of Turing machine.

Now, having  $A \subset \mathcal{A}^*$ ,  $P \in \underline{\text{Pred}}_1$ ,  $x \in A$ ,  $\rho(A, P, x)$  may be defined as the number of steps (actions) taken by a given  $\mathcal{U}$ -Turing machine  $T_P$  during the computation of the truth value of the formula  $P(x)$  (of the characteristic function  $\chi_{A|P}(x)$ , in the set-theoretic setting). If  $T_P$  does not terminate its work for some  $x \in A$ ,  $\rho(A, P, x) = \infty$ . We shall write  $\rho_{\mathcal{U}}(A, P, x)$ , if necessary or appropriate to express explicitly the role of  $\mathcal{U}$  in what follows.

Another example of numerical reference systems will be perhaps more intuitive.

Let  $X, A \subset X$ ,  $\mathcal{L}$ , and  $\underline{\text{Pred}}$  be as in the general definition of reference systems above, let  $\underline{\text{Pred}}_0$  be a finite or infinite subset of  $\underline{\text{Pred}}$ . As a rule, predicates in  $\underline{\text{Pred}}_0$  are supposed to be logically independent and atomic pieces of knowledge from which all other predicates under consideration are built through logical connectives and quantifiers (but quantifiers will not be taken into account in what follows). Atomic formulas generated by predicates from  $\underline{\text{Pred}}_0$  together with their negations will be called *literals*. If  $R_{i_1}(x), R_{i_2}(x), \dots, R_{i_{n(i)}}(x)$  is a finite sequence of literals with the same indeterminate  $x$  and with no predicate occurring twice or more times, then

$$\begin{aligned} \alpha_i(x) &=_{\text{df}} \bigwedge_{j=1}^{n(i)} R_{i_j}(x) \\ &=_{\text{df}} R_{i_1}(x) \text{ and } R_{i_2} \text{ and } \dots \text{ and } R_{i_{n(i)}}(x) \end{aligned} \quad (23)$$

is the so called *elementary conjunction* (over  $\underline{\text{Pred}}_0$ ). *Normal predicate* (over  $\underline{\text{Pred}}_0$ ) is defined by a disjunction of elementary conjunctions, i.e., by a formula  $S(x)$  of the form

$$\begin{aligned} S(x) &=_{\text{df}} \bigvee_{i=1}^K \alpha_i(x) =_{\text{df}} \alpha_1(x) \text{ vel } \alpha_2(x) \text{ vel } \dots \text{ vel } \alpha_K(x) = \\ &= \bigvee_{i=1}^K \bigwedge_{j=1}^{n(i)} R_{i_j}(x); \end{aligned} \quad (24)$$

the set of normal predicates over  $\underline{\text{Pred}}_0$  is denoted by  $\underline{\text{Pred}}_1$ . As can be easily seen, every predicate from  $\underline{\text{Pred}}_1$ , formed by propositional connectives from predicates in  $\underline{\text{Pred}}_0$ , is logically equivalent to a predicate from  $\underline{\text{Pred}}_1$ .

Now, consider a testing oracle or device which is able, at least for some  $x \in A$  and  $P \in \underline{\text{Pred}}_0$ , to verify whether  $P(x)$  holds or not. Let  $c(P, x)$  denote the quantitative cost of such a verification, e.g., time or other demands or expenses and suppose that  $c(P, x) = \infty$  iff the device is not able to test  $P(x)$  ( $c(P, x)$  is non-negative in all cases). The testing device is able to prove the validity of  $S(x)$  defined by (24) iff there is at least one  $\alpha_i(x)$ ,  $i \leq K$  such that  $R_{i_j}(x)$  holds for each  $j \leq n(i)$ . Set

$$m_i(S, x) = \min \{k : k \leq n(i), R_{i_k}(x) \text{ is false}\}, \quad (25)$$

if such an index  $k$  exists,

$$m_i(S, x) = n(i) \quad (26)$$

otherwise. Set, moreover,

$$q_i(S, x) = \sum_{j=1}^{m_i(S, x)} c(\tilde{R}_{i_j}(x)), \quad (27)$$

where  $\tilde{R}_{i_j}$  is the predicate from  $\underline{\text{Pred}}_0$  occurring in  $R_{i_j}(x)$ . Now, define

$$M_1(S, x) = \max \{q_i(S, x) : i \leq K\}, \quad (28)$$

$$M_2(S, x) = \sum_{i=1}^{K_1} q_i(S, x) \quad (29)$$

if

$$K_1 = \min \left\{ L : L \leq K, m_L(S, x) = n(L), R_{L_{n(L)}} \text{ holds} \right\} \quad (30)$$

is defined, and

$$M_2(S, x) = \sum_{i=1}^K q_i(S, x) \quad (31)$$

otherwise.

The intuition behind is simple, but let us introduce it explicitly. Let the testing device verify  $\alpha_i(x)$  defined by (23), so, first of all,  $R_{i_1}(x)$  is tested, where  $R_{i_1}(x)$  is either  $P(x)$  or non  $P(x)$  for some  $P \in \underline{\text{Pred}}_0$ . If  $R_{i_1}(x)$  does not hold, the testing device proclaims  $\alpha_i(x)$  to be false and terminates its work (invalidity of one component trivially implies the invalidity of all the conjunction). If  $R_{i_1}(x)$  holds, the device tests  $R_{i_2}(x)$  and either terminates its work (if  $R_{i_2}(x)$  is false or if it is the last component in  $\alpha_i(x)$ ), or goes on with  $R_{i_3}(x)$ , if  $R_{i_2}(x)$  holds and  $R_{i_3}(x)$  exists. Hence,  $m_i(S, x)$ , which can be also write as  $m(\alpha_i(x))$ , is nothing else than the index of the last component of  $\alpha_i(x)$  which is tested before the validity of  $\alpha_i(x)$  can be proved, and  $q_i(S, x)$  is the sum of expenses connected with the verification of the actually tested members  $R_{i_1}(x), R_{i_2}(x), \dots, R_{i_{m_i(S, x)}}(x)$  of  $\alpha_i(x)$ , so that  $q_i(S, x)$  can be taken as the cost of verifying of  $\alpha_i(x)$ .

Suppose that the elementary conjunctions in  $S$  are tested in parallel and, having verified all of them, a supervisor decides, whether there is a valid elementary conjunction among  $\alpha_1(x), \alpha_2(x), \dots, \alpha_K(x)$ . If it is the case,  $S$  is proclaimed to be valid, in the opposite case  $S$  is proclaimed to be false. If, moreover, the expenses connected with testing of  $\alpha_i(x)$  are taken as time demands (time computational complexity), and the expenses connected with the inspection and cumulation of particular outputs by the supervisor are neglected, then the maximum value of  $q_i(S, x)$ , i. e.  $M_1(S, x)$ , expresses the expenses (time computational complexity) corresponding to the verification of  $S$  by the parallel testing device in question. The simplifications just introduced, concerning the neglect of the expenses connected with the supervisor activity, are the common ones in the theory of nondeterministic algorithms.

On the other hand side,  $M_2(S, x)$  reflects the case when particular  $\alpha_i(x)$ 's in  $S$  are checked sequentially, first,  $\alpha_1(x)$ , if it is not true, then  $\alpha_2(x)$ , and so on till such an  $\alpha_{K_1}(x)$ , which is valid (true), i. e. all  $R_{K_1, j}(x)$ ,  $j \leq n(K_1)$ , are valid. In this case the testing device terminates its work and proclaim  $S$  to hold (the validity of one component trivially implies the validity of all the disjunction). The value  $M_2(S, x)$  then cumulates the total expenses connected with the testing of all elementary conjunctions till the  $K_1$ -th. If this is not the case, i. e., if each  $\alpha_i(x)$  can be disproved through finding some  $R_{i_j}(x)$  to be false, then evidently all the elementary conjunctions  $\alpha_1(x), \alpha_2(x), \dots, \alpha_K(x)$  must be checked in order to find that  $S$  does not hold, hence, in this case the total expenses are given by (31).

So, combining all our reasonings and going back to the problem of an appropriate definition of a numerical reference system for  $X$ ,  $A \subset X$ , and  $\underline{\text{Pred}}_1 \subset \underline{\text{Pred}}$  as above,

we may set, quite reasonably, for  $S \in \underline{\text{Pred}}$ , and  $x \in A$

$$\rho(A, S, x) = M_i(S, x), \quad (32)$$

choosing  $i = 1$  or  $i = 2$  according to the sequential or parallel interpretation or implementation of the corresponding verification algorithm.

Keeping in mind the fact that reference systems play only auxiliary and intermediate role on the way leading to the rough sets, the time and space devoted to these systems in Chapters 3 and 4 may seem to be too large. However, we believe that at least an informal idea concerning some particular reference systems enables, to the reader, to understand better and more easily the explanations and reasonings concerning the rough sets defined through such reference systems leaving aside, for this moment, very interesting philosophical and set-theoretical problems connected with the relativization of axiom of extensionality enabled by our approach.

#### 5. CONSISTENT AND OPTIMAL $b$ -ROUGH SETS

**Definition 3.** Let  $X, A \subset X, \mathcal{L}, \underline{\text{Pred}}, \rho$ , and  $\langle B, \preceq \rangle$  be as in Chapter 2, let  $V$  be a (classical) subset of  $A$ . Now,  $b$ -rough set consistent with  $V$  is a pair  $\langle V_*, V^* \rangle$  of sets such that  $V_*, V^* \in \mathcal{F}(b, A, \rho)$ ,  $V_* \subset V \subset V^*$ , where

$$\mathcal{F}(b, A, \rho) = \{C \subset A : (\exists P \in \underline{\text{Pred}})(C = A \mid P, (\forall x \in A)(\rho(A, P, x) \preceq b))\}. \quad (33)$$

A  $b$ -rough set  $\langle V_*, V^* \rangle$  consistent with  $V$  is called *optimal* (for  $V$ ), if for each  $b$ -rough set  $\langle V_{**}, V^{**} \rangle$  consistent with  $V$  the inclusions  $V_{**} \subset V_*$  and  $V^* \subset V^{**}$  hold (informally, if  $\langle V_*, V^* \rangle$  is “the best” approximation or description of  $V$  achievable by  $b$ -subsets of  $A$ ).

In the classical approach to rough sets, briefly reviewed in Chapter 1, the given subset  $V \subset A \subset X$  and the given equivalence relation  $\approx$  on  $X$  uniquely define the sets  $V_*, V^*$ , and the corresponding rough set. In the alternative setting presented here,  $V, \rho, \langle B, \preceq \rangle$ , and  $b \in B$  need not define only one  $b$ -rough set consistent with  $V$ . Hence, the existence of the  $b$ -rough set optimal for  $V \subset A$  is not trivial; in what follows, we shall present an intuitively interpretable and justifiable system of conditions under which the existence of optimal  $b$ -rough sets is evident.

**Definition 4.** The language  $\mathcal{L}$  is called *complete*, if there exist, for each nonempty set  $\mathcal{P} \subset \underline{\text{Pred}}$  of unary predicates of  $\mathcal{L}$  two unary predicates  $Q_*(\mathcal{P}) \in \underline{\text{Pred}}$ ,  $Q^*(\mathcal{P}) \in \underline{\text{Pred}}$ , such that

$$(\forall x \in A)[Q_*(\mathcal{P})(x) \text{ iff } (\exists P \in \mathcal{P})(P(x))], \quad (34)$$

and

$$(\forall x \in A)[Q^*(\mathcal{P})(x) \text{ iff } (\forall P \in \mathcal{P})(P(x))], \quad (35)$$

where  $P \text{ iff } Q$  abbreviates  $(P \Rightarrow Q)$  and  $(Q \Rightarrow P)$ .

The reference system  $\rho$  over a complete language  $\mathcal{L}$  is called *parallel*, if for each nonempty  $\mathcal{P} \subset \underline{\text{Pred}}$ ,

$$\rho(A, Q_*(\mathcal{P}), x) \preceq \sup \{ \rho(A, P, x) : P \in \mathcal{P} \}, \quad (36)$$

and

$$\rho(A, Q^*(\mathcal{P}), x) \preceq \sup \{ \rho(A, P, x) : P \in \mathcal{P} \}. \quad (37)$$

Here, inf and sup extend the definitions of  $\wedge$  and  $\vee$ , given by (6) and (7), in such a way that

$$\underline{\text{inf}} \{ \rho(A, P, x) : P \in \mathcal{P} \} \preceq \rho(A, P, x) \preceq \sup \{ \rho(A, P, x) : P \in \mathcal{P} \} \quad (38)$$

for all  $P \in \mathcal{P}$ , moreover, if

$$y \preceq \rho(A, P, x) \preceq z \quad (39)$$

for all  $P \in \mathcal{P}$ , then

$$y \preceq \underline{\text{inf}} \{ \rho(A, P, x) : P \in \mathcal{P} \} \quad (40)$$

and

$$\sup \{ \rho(A, P, x) : P \in \mathcal{P} \} \preceq z. \quad (41)$$

The adjective “parallel” connected with certain reference system is almost self-evident and reflects the ability of such systems to test, given  $x \in A$ , simultaneously, hence, in parallel, for all  $P \in \mathcal{P}$ , whether  $P(x)$  holds or not. Within the (time) complexity sup  $\{ \rho(A, P, x) : P \in \mathcal{P} \}$  this question is answered for all  $P \in \mathcal{P}$ . Then, a supervisor may ask either whether there is  $P \in \mathcal{P}$  such that  $P(x)$  holds, or whether  $P(x)$  holds for each  $P \in \mathcal{P}$ , in order to decide, whether  $Q_*(\mathcal{P})(x)$  or  $Q^*(\mathcal{P})(x)$  hold. The expenses and demands of all kinds connected with the supervisor activity are neglected according to the classical paradigma of nondeterministic and parallel algorithms. Let us remark that it is not possible to replace sup by inf in (36). Or, having do so, the reference system would stop its parallel verifying of  $P(x)$  for different  $P \in \mathcal{P}$ , having decided  $P_0(x)$  for some  $P_0 \in \mathcal{P}$ . However, if this decision is the negative one, i. e. if non  $P_0(x)$  holds, we will not be able to say, whether  $Q_*(\mathcal{P})(x)$  holds or not (perhaps  $P_1(x)$  holds for another  $P_1 \in \mathcal{P}$ ).

**Fact 3.** If  $\rho$  is a parallel reference system over a complete language  $\mathcal{L}$ , then there exists, for each  $b \in B$  such that  $\emptyset, A \in \mathcal{F}(b, A, \rho)$  and each  $V \subset A$ , the optimum  $b$ -rough set for  $V$ .

*Proof.* Let  $V \subset A$ , let  $b \in B$  be such that  $\emptyset, A \in \mathcal{F}(b, A, \rho)$ , set

$$\mathcal{F}(b, A, \rho)_* V = \{ C : C \in \mathcal{F}(b, A, \rho), C \subset V \}, \quad (42)$$

$$\mathcal{F}(b, A, \rho)^* V = \{ C : C \in \mathcal{F}(b, A, \rho), C \supset V \}. \quad (43)$$

Let  $\mathcal{P}_{*V}$  ( $\mathcal{P}_V^*$  resp.) denote the subset of those unary predicates from Pred which define sets from  $\mathcal{F}(b, A, \rho)_*V$  (from  $\mathcal{F}(b, A, \rho)^*V$ , resp.), in symbols

$$\mathcal{P}_{*V} = \{P \in \text{Pred} : (\exists C \in \mathcal{F}(b, A, \rho))(C = A \mid P \subset V)\}, \quad (44)$$

$$\mathcal{P}_V^* = \{P \in \text{Pred} : (\exists C \in \mathcal{F}(b, A, \rho))(C = A \mid P \supset V)\}. \quad (45)$$

Due to the assumptions,  $\emptyset \in \mathcal{F}(b, A, \rho)_*V$  and  $A \in \mathcal{F}(b, A, \rho)^*V$ , hence,  $\mathcal{P}_{*V} \neq \emptyset$  and  $\mathcal{P}_V^* \neq \emptyset$ . Set

$$V_* = \cup \{C : C \in \mathcal{F}(b, A, \rho)_*V\}, \quad (46)$$

$$V^* = \cap \{C : C \in \mathcal{F}(b, A, \rho)^*V\}. \quad (47)$$

Evidently, for each  $b$ -rough set  $\langle V_*, V^{**} \rangle$  consistent with  $V, V_* \subset V_*$  and  $V^* \subset V^{**}$ , the only one we have to prove is that  $\langle V_*, V^* \rangle$  is a  $b$ -rough set.  $\mathcal{L}$  is supposed to be a complete language, so that we may write

$$V_* = A \mid [Q_*(\mathcal{P}_{*V})], \quad (48)$$

$$V^* = A \mid [Q^*(\mathcal{P}_V^*)], \quad (49)$$

where  $Q_*(\mathcal{P}_{*V})$  and  $Q^*(\mathcal{P}_V^*)$  are in Pred. As  $\rho(A, P, x) \preceq b$  for each  $x \in A$ , (36), (37) and (41) imply

$$\rho(A, Q_*(\mathcal{P}_{*V}), x) \preceq b, \quad (50)$$

$$\rho(A, Q^*(\mathcal{P}_V^*), x) \preceq b, \quad (51)$$

for all  $x \in A$ . Hence,  $\langle V_*, V^* \rangle$  is a  $b$ -rough set consistent with  $V$ , and the assertion is proved.  $\square$

## 6. NUMERICALLY QUANTIFIED ROUGH SETS AND DEMPSTER-SHAFER THEORY

Let us consider, in this chapter, a rough set  $\langle V_*, V^* \rangle$  in the basic space  $X$  defined by, or associated with, a ‘‘classical’’ subset  $V \subset X$ . This can be done either through an equivalence relation on  $X$ , or through a predicate attributed to elements of  $X$ , supposing that the ineffective or at least infeasible nature of this predicate does not enable to decide effectively about its validity for each  $x \in X$ . In every case, the inclusion  $V_* \subset V \subset V^*$  evokes the idea to take  $\langle V_*, V^* \rangle$  as a ‘‘confidence interval’’ for  $V$ ; another immediately arising idea is to quantify somehow, using real numbers, the ‘‘tightness’’ of the ‘‘interval’’  $\langle V_*, V^* \rangle$ . This is meant in the same sense as that in which the length or, more generally, a Lebesgue or Lebesgue–Stieltjes measure of an interval on the real line quantitatively expresses the degree or quality of setting of a point just known to be covered by this interval.

Even when intentionally limiting ourselves to such numerical quantifications of rough sets which are compatible with an appropriate probabilistic interpretation, we may arrive

at such quantifications using different ways. Let us consider the two ones closely connected with the two approaches to the definition of rough sets as explained and discussed above.

Let  $\langle V_*, V^* \rangle$  be defined, using (1) and (2), by  $V \subset X$  and by an equivalence relation  $\approx$  defined on  $X$ . In the most simple case, when  $Pr$  is a probability measure defined for each subset of the space  $X_0 = X/\approx$  of equivalence classes generated by  $X$  and  $\approx$ , we may simply ascribe to the rough set  $\langle V_*, V^* \rangle$  the pair (or interval) of real numbers  $\langle Pr(V_*), Pr(V^*) \rangle$ , evidently  $Pr(V_*) \leq Pr(V^*)$ . Let us recall, that  $Pr$  is a probability (measure) on the set of all subsets of  $X_0$ , if  $0 \leq Pr(V_0) \leq 1$  for each  $V_0 \subset X_0$ ,  $Pr(\emptyset) = 0$ ,  $Pr(X_0) = 1$  ( $\emptyset$  is the empty set), and  $Pr(\bigcup_{i=1}^{\infty} V_i) = \sum_{i=1}^{\infty} Pr(V_i)$  for each infinite sequence  $V_1, V_2, \dots$  of mutually disjoint subsets of  $X_0$ . It is worth an explicit noting that the value  $Pr(V)$  need not be defined (if  $V_* \neq V^*$ ), moreover, the probability measure  $Pr$  need not be consistently extendable in such a way that  $Pr(V)$  were defined (there are subsets of, say, the unit interval of real numbers to which cannot be ascribed a probability measure consistent with this one ascribing to each interval its length). Hence, the pair or interval  $\langle Pr(V_*), Pr(V^*) \rangle$  must be taken as a numerical characteristic or confidence interval for the rough set  $\langle V_*, V^* \rangle$ , not for a defined but perhaps unknown value  $Pr(V)$ .

This definition can be easily seen to copy the definitions of inner and outer measures in the measure theory (cf., e.g. Halmos [3]). Going on in this spirit we may generalize this definition as follows. Let  $\mathcal{S}$  be a system of subsets of  $X_0$  for which the probability  $Pr$  is defined and suppose, for the sake of simplicity, that  $\emptyset \in \mathcal{S}$ ,  $X_0 \in \mathcal{S}$ . The axiom of  $\sigma$ -additivity reduces to those infinite sequences of mutually disjoint subsets from  $\mathcal{S}$ , the union of which is also in  $\mathcal{S}$ . Now, the rough set  $\langle V_*, V^* \rangle$  can be quantified by the pair  $\langle Pr_*(V_*, \mathcal{S}), Pr^*(V^*, \mathcal{S}) \rangle$  of real numbers defined by

$$\begin{aligned} Pr_*(V_*, \mathcal{S}) &= \\ &= \sup \left\{ \sum_{i=1}^{\infty} Pr(V_i) : V_i \in \mathcal{S}, V_i \subset V_*, i \neq j \Rightarrow V_i \cap V_j = \emptyset, i, j = 1, 2, \dots \right\}, \end{aligned} \quad (52)$$

$$Pr^*(V^*, \mathcal{S}) = 1 - Pr_*(A - V^*, \mathcal{S}). \quad (53)$$

The explicit condition that  $\mathcal{S}$  contains only subsets of  $X_0$  can be immediately generalized by saying that  $\mathcal{S}$  contains some subsets of  $X$ ; their definability in  $X_0$  may be a part of the definition of  $\mathcal{S}$ . Of course, some variants of (52) and (53) may be also considered.

At the very beginning of this chapter we considered the rough set  $\langle V_*, V^* \rangle$  generated by a set  $V \subset X$ . This rough set admits an immediate and trivial re-interpretations in the logical terms, as if  $x \in V_*$ , then evidently  $x \in V$ , on the other side,  $x \in X - V^*$  implies  $x \in X - V$ . Hence, the validity of  $x \in V_*$  can be taken as a *sufficient*, and the validity of  $x \in V^*$  as a *necessary* condition for  $x \in V$  to hold. Because of our limited abilities either to discern different points of the universe  $X$ , or to evaluate, given  $x \in X$ , the truthvalue of the predicate  $P$  which defined  $V$  in  $X$ , we are often not able to give a necessary and (simultaneously) sufficient condition for  $x \in V$ , so that this interpretation

of rough sets quite agrees with the original one taking rough sets as a set-theoretical tool to express, in formalized terms, our gnoseological limitations. On the other side, taking the numerical value  $Pr(V_*)$  as the probability, or at least as a probabilistic numerical evaluation, of a sufficient condition for the validity of the predicate  $x \in V$ , taken as a decision or as a conclusion resulting from some considerations, we have arrived very close to the ideas and ways of reasoning presented and formalized in the so called Dempster-Shafer theory, very popular in our days (cf., e.g. [15], [16], [17]). The closeness is so tight that it deserves a more detailed investigation. Attempting to do so in the rest of this chapter, we begin with a short description of the basic ideas of Dempster-Shafer approach.

For the sake of simplicity let us consider the finite case. Let  $\mathcal{P} = \{p_1, p_2, \dots, p_m\}$ ,  $\mathcal{Q} = \{q_1, q_2, \dots, q_n\}$  be two sequences, not necessarily disjoint, of elementary formulas or, formally, propositional indeterminates. These elementary formulas, or atomic pieces of information concerning the problem and situation in question, are supposed to be logically independent, so that for each sequence  $r_1, r_2, \dots, r_k$  of elements from  $\{p_i, \neg p_i, q_j, \neg q_j, i \leq m, j \leq n\}$ , with each  $p_i$  and  $q_j$  occurring at most once, the conjunction  $r_1 \wedge r_2 \wedge \dots \wedge r_k$  is logically consistent. Here  $p_i, q_j, \neg p_i, \neg q_j$  in common are called *literals*,  $r_i \wedge r_j$  stands for  $r_i$  and  $r_j$ ,  $\neg p_i$  stands for non  $p_i$ . Set

$$U(\mathcal{P}) = \{r_1 \wedge r_2 \wedge \dots \wedge r_m : r_i \in \{p_i, \neg p_i\}, i \leq m\}, \quad (54)$$

$$U(\mathcal{Q}) = \{r_1 \wedge r_2 \wedge \dots \wedge r_n : r_i \in \{q_j, \neg q_j\}, j \leq n\}. \quad (55)$$

Elements of  $U(\mathcal{P})$  can be called *elementary* or *atomic states* (states of world, worlds), elements of  $U(\mathcal{Q})$  can be called *elementary* or *atomic hypotheses*. Formulas of the propositional language  $\mathcal{L}(\mathcal{P})$ , generated in the usual way by  $p_1, p_2, \dots, p_m, \wedge, \neg$ , and possibly other additionally defined propositional connectives, are called *conditions*. They can be easily and unambiguously identified with subsets of  $U(\mathcal{P})$ , ascribing to each formula of  $\mathcal{L}(\mathcal{P})$  the set of elementary conjunctions of literals, occurring in the disjunctive normal form of this formula. Formulas of the propositional language  $\mathcal{L}(\mathcal{Q})$ , generated analogously by  $q_1, q_2, \dots, q_n$ , are called *hypotheses* and are identified with subsets of  $U(\mathcal{Q})$ . *Evidence* are formulas of the propositional language  $\mathcal{L}(\mathcal{P} \cup \mathcal{Q})$ , generated by all the propositional indeterminates occurring either in  $\mathcal{P}$ , or in  $\mathcal{Q}$ , again, evidence can be identified with subsets of  $U(\mathcal{P} \cup \mathcal{Q})$ . Having arrived at this stage of explanation, we could work, in what follows, with abstract sets  $\mathcal{S} (= U(\mathcal{P}))$  of states, and  $\mathcal{H} (= U(\mathcal{Q}))$  of elementary hypotheses, taking conditions as subsets of  $\mathcal{S}$ , hypotheses as subsets of  $\mathcal{H}$ , and evidence as subsets of the Cartesian product  $\mathcal{S} \times \mathcal{H}$ . However, because of having at hand a more transparent intuitive image, we shall go on with the previous interpretation based on the meta-language of the propositional language.

Consider a fixed finite set of pieces of evidence  $E_1, E_2, \dots, E_k \in \mathcal{L}(\mathcal{P} \cup \mathcal{Q})$ . These formulas describe, even if only partially, some relations among the states of world (definable in the language  $\mathcal{L}(\mathcal{P})$ ), and certain hypotheses (definable in the language  $\mathcal{L}(\mathcal{Q})$ ). Hence, given a formula  $B$  of  $\mathcal{L}(\mathcal{Q})$  and denoting by  $\vdash$  the usual meta-theoretical relation

of logical consequence, we may define the set  $\mathcal{C}(B)$  of formulas of  $\mathcal{L}(\mathcal{P})$  in this way:

$$\mathcal{C}(B) = \{A : A \in \mathcal{L}(\mathcal{P}); A, E_1, \dots, E_k \vdash B\}, \quad (56)$$

denoting by  $\mathcal{C}_0(B)$  the subset of consistent formulas from  $\mathcal{C}(B)$ . An alternative definition can be given in the set-theoretic language. Or, denoting by  $S_A$  the subset of  $\mathcal{S}$  (of  $U(\mathcal{P})$ ), in the particular interpretation corresponding to  $A \in \mathcal{L}(\mathcal{P})$ , and denoting, for  $B \in \mathcal{L}(\mathcal{Q})$ , by  $S_B$  the corresponding subset of  $\mathcal{H}$  or  $U(\mathcal{Q})$ , (56) turns into

$$\begin{aligned} \mathcal{S}_C(B) &= \{S_A : A \in \mathcal{C}(B)\} = \\ &= \{S \subset \mathcal{S} : \vdash (\forall s \in \mathcal{S}, t \in \mathcal{H}) [(s \in S) \wedge E_1 \wedge \dots \wedge E_k \rightarrow (t \in B)]\}. \end{aligned} \quad (57)$$

Now, let  $\pi$  be a probability distribution over consistent formulas of the language  $\mathcal{L}(\mathcal{P})$  with  $\pi^*$  denoting the corresponding uniquely defined probability distribution over nonempty subsets of  $U(\mathcal{P})$ . Hence,  $\pi : \mathcal{L}(\mathcal{P}) \rightarrow \{0, 1\}$ , if  $A$  is inconsistent, then  $\pi(A) = 0$  and  $\sum_{A \in \mathcal{L}(\mathcal{P})} \pi(A) = 1$ . Analogously,  $\pi^* : 2^{U(\mathcal{P})} \rightarrow \{0, 1\}$ ,  $\pi^*(\emptyset) = 0$ , and  $\sum_{A \subset U(\mathcal{P})} \pi^*(A) = 1$ , where  $\pi(A) = \pi^*(S_A)$ . For  $B \in \mathcal{L}(\mathcal{Q})$ , *believability*  $Bel_\pi(B)$  of  $B$ , or  $Bel_\pi(S_B)$  in the set language, is defined by

$$Bel_\pi(B) = \sum_{A \in \mathcal{C}_0(B)} \pi(A), \quad (58)$$

the summation can be taken over whole  $\mathcal{C}(B)$  due to the assumptions.

*Plausibility*  $Pl_\pi(B)$  of  $B$  is defined by

$$Pl_\pi(B) = 1 - Bel_\pi(\neg B). \quad (59)$$

The index  $\pi$  will be omitted supposing the corresponding probability distribution is fixed or unambiguously given by the context.

Consider the set-theoretic definition (57). Clearly, if  $S_1, S_2 \in \mathcal{S}_C(B)$ , then also  $S_1 \cup S_2 \in \mathcal{S}_C(B)$ , if  $S_1 \subset S_2 \in \mathcal{S}_C(B)$ , then  $S_1 \in \mathcal{S}_C(B)$  as well. Hence, given  $B \in \mathcal{L}(\mathcal{Q})$  or the corresponding set  $T_B \in U(\mathcal{Q})$ , there exists unique  $S_B \subset U(\mathcal{P})$  such that  $\mathcal{S}_C(B) = \{A : A \subset S_B\}$ . Let  $\hat{S}$  be a random variable defined on an abstract probability space  $(\Omega, \mathcal{S}, P)$ , taking its values in the space  $2^{U(\mathcal{P})}$  of all subsets of  $U(\mathcal{P})$  and such that

$$P(\{\omega : \omega \in \Omega, \hat{S}(\omega) = A\}) = \pi(A) \quad (60)$$

with  $\pi$  defined as above. As can be easily seen, (58) is equivalent to

$$Bel_\pi(B) = P(\{\omega : \omega \in \Omega, \hat{S}(\omega) \subset S_B\}). \quad (61)$$

Dempster-Shafer theory can be translated or interpreted into the terms of rough sets and their numerical quantifications as follows. Let us suppose that there exists, at an objective level, a mapping  $F_0$  ascribing to each elementary state of world  $s \in \mathcal{S} = U(\mathcal{P})$  the corresponding elementary hypothesis  $F_0(s) = t \in \mathcal{H} = U(\mathcal{Q})$ ; this elementary

hypothesis is supposed to be the only one valid under the condition that  $s$  is the actual state of world. However, the mapping  $F_0$  is not known completely, the only partial knowledge about  $F_0$  is given through the evidence  $E$  taken as a subset of the Cartesian product  $\mathcal{S} \times \mathcal{H}$ . Namely, we know that, for each  $s \in \mathcal{S}$ ,  $\langle s, F_0(s) \rangle \in E$ .

In the spirit of the Dempster–Shafer approach, the only we know (or suppose to know) about the actual state of world  $s_0$  is that it lies in the subset  $\tilde{S}(\omega)$  of  $\mathcal{S}$  sampled at random by a random variable  $\tilde{S}$  taking a probability space  $(\Omega, \mathcal{A}, Pr)$  into  $2^{\mathcal{S}} - \{\emptyset\}$ . This uncertainty concerning the state  $s_0$  and, consequently, also the corresponding elementary hypothesis  $F_0(s_0)$ , can be formally expressed by introducing two random variables  $X, Y$ , both defined on  $(\Omega, \mathcal{A}, Pr)$  and taking their values in  $\mathcal{S}$  (for  $X$ ) or in  $\mathcal{H}$  (for  $Y$ ), even if these values are not completely accessible for the observer. Hence, the actual state of world is  $X(\omega)$ , the corresponding elementary hypothesis is  $Y(\omega) = F_0(X(\omega))$ , and the only we know is that  $\langle X(\omega), Y(\omega) \rangle \in E$ .

Take, for a while, the abstract set  $\Omega$ , supporting the probability space in question, as a set  $\mathcal{S}_0$  of elementary states and suppose that we are able to prove or disprove all assertions of the form  $Y(\omega) \in B$ ,  $B \in \mathcal{H}$ , just given  $\omega \in \Omega$  and knowing that  $\langle X(\omega), Y(\omega) \rangle \in E$ . All elements of  $\Omega$  are supposed to be completely distinguishable. Then the set  $A = A(B, E) \subset \Omega$ , defined by

$$A = \{\omega : \omega \in \Omega, \langle X(\omega), Y(\omega) \rangle \in E, Y(\omega) \in B\} \quad (62)$$

coincides with the set  $S_B$  defined by the Dempster–Shafer theory when  $B$  is replaced by  $\mathcal{S}_0$ . Similarly, omitting the symbols  $\dots \omega \in \Omega \dots$  in what follows, the set

$$A_1 = \{\omega : \langle X(\omega), Y(\omega) \rangle \in E, Y(\omega) \in \mathcal{H} - B\} \quad (63)$$

can be taken as  $S_{\mathcal{H}-B}$ .  $S_B$  and  $S_{\mathcal{H}-B}$  are disjoint and

$$S_B \cup S_{\mathcal{H}-B} = A \cup A_1 = \{\omega : \langle X(\omega), Y(\omega) \rangle \in E\}. \quad (64)$$

As a matter of fact, however, elements of  $\Omega$  are distinguishable from each other only partially, through different values of observable random variables taken for different elements from  $\Omega$ . So we may introduce an equivalence relation  $\approx_X$  defined on  $\Omega$  by the random variable  $X$  as follows: for  $\omega_1, \omega_2 \in \Omega$ ,  $\omega_1 \approx_X \omega_2$  iff  $X(\omega_1) = X(\omega_2)$ . In fact,  $\approx_X$  evidently is an equivalence relation. Now, let  $\langle A_X, A^X \rangle$  be the rough set generated by  $A$  and  $\approx_X$  in  $\Omega$ , hence,

$$\begin{aligned} A_X &= \{\omega : \langle X(\omega), Y(\omega) \rangle \in E, Y(\omega) \in B, (\forall \omega_1 \in \Omega) [(X(\omega_1) \\ &= X(\omega)) \Rightarrow \langle X(\omega_1), Y(\omega_1) \rangle \in E \text{ et } Y(\omega_1) \in B]\}, \end{aligned} \quad (65)$$

$$\begin{aligned} A^X &= \Omega - (\Omega - A)_X = \{\omega : (\exists \omega_1 \in \Omega) [(X(\omega) = X(\omega_1)) \text{ et} \\ &\langle \langle X(\omega_1), Y(\omega_1) \rangle \in E \rangle \text{ et } \langle Y(\omega_1) \in B \rangle]\}. \end{aligned} \quad (66)$$

Being a random variable, the mapping  $X$  is measurable, and as  $\mathcal{S}$  is supposed to be finite we obtain that for each  $s \in \mathcal{S}$

$$\{\omega : \omega \in \Omega, X(\omega) = s\} \in \mathcal{A}. \quad (67)$$

Moreover, both  $A_X$  and  $A^X$  are finite unions of sets of this kind, so that  $A_X, A^X \in \mathcal{A}$  and  $Pr(A_X), Pr(A^X)$  are defined. Evidently,  $A_X \subset A^X$  and  $Pr(A_X) \leq Pr(A^X)$ , so that the pair  $(Pr(A_X), Pr(A^X))$  can be taken as a probabilistic numerical quantification of the rough set  $\langle A_X, A^X \rangle$ .

This rough set  $\langle A_X, A^X \rangle$  generates another rough set  $\langle S_{X,A}, S^{X,A} \rangle$  in  $\mathcal{S}$  as follows. Denote, for each  $s \in \mathcal{S}$ , by  $X^{-1}(s)$  the subset of  $\Omega$  defined by (67) and define

$$S_{X,A} = X(A_X) = \{s : s \in \mathcal{S}, X^{-1}(s) \subset A_X\} = \bigcup_{\omega \in A_X} \{X(\omega)\}, \quad (68)$$

$$\begin{aligned} S^{X,A} &= \mathcal{S} - X(\Omega - A^X) = \mathcal{S} - \{s : s \in \mathcal{S}, X^{-1}(s) \subset \Omega - A^X\} = \\ &= \{s : X^{-1}(s) \subset A^X\}, \end{aligned} \quad (69)$$

as for each  $s \in \mathcal{S}$  either  $X^{-1}(s) \subset A^X$  or  $X^{-1}(s) \subset \Omega - A^X$ . The inclusion  $A_X \subset A^X$  immediately implies that  $S_{X,A} \subset S^{X,A}$ , so that the pair  $\langle S_{X,A}, S^{X,A} \rangle$  can be taken as a rough set in  $\mathcal{S}$ . Intuitively, if  $X(\omega) = s \in S_{X,A}$ , we are able to prove that  $Y(\omega) \in B$  using this deduction:  $X(\omega) = s \in S_{X,A}$  yields  $X^{-1}(s) \subset A_X$ , the "actual state of world"  $\omega_0$  is in  $X^{-1}(s)$ , hence,  $\omega_0 \in A_X \subset A$  so that  $Y(\omega_0) \in B$  due to (62). Similarly, if  $X(\omega) \in \mathcal{S} - S^{X,A}$ , we are able to prove that  $Y(\omega) \in \mathcal{H} - B$ , as  $\langle X(\omega), Y(\omega) \rangle \in E$  is supposed to hold in every case.

Worth an explicit mentioning is the fact that, in general, the rough set  $\langle S_{X,A}, S^{X,A} \rangle$  is not generated by a set  $S_0 \subset \mathcal{S}$  satisfying

$$\begin{aligned} (\forall s \in \mathcal{S}) \{ \{s \in S_0\} \underline{\text{et}} (X(\omega) = s) \underline{\text{et}} ((X(\omega), Y(\omega)) \in E) \Rightarrow \\ (Y(\omega) \in B) \} \underline{\text{et}} \{ \{s \in \mathcal{S} - S_0\} \underline{\text{et}} (X(\omega) = s) \underline{\text{et}} ((X(\omega), Y(\omega)) \in E) \\ \rightarrow (Y(\omega) \in \mathcal{H} - B) \} \}, \end{aligned} \quad (70)$$

as such a set need not exist neither in the classical sense.

We may arrive to the original rough set  $\langle A_X, A^X \rangle$  also by the alternative way consisting in an appropriate quantification of the difficulties connected with decision making for a predicate defining  $A$  as a subset of  $\Omega$ . This role can evidently play the predicate

$$P_1(\cdot) = ((X(\cdot), Y(\cdot)) \in E) \underline{\text{et}} (Y(\cdot) \in B) \quad (71)$$

used in (62).

Let  $B = (b_1 \preceq b_2 \preceq b_3, \text{non } (b_3 \preceq b_2))$  be a partially (in fact, linearly) ordered set, let  $\rho$  be a monotonous reference system taking its values in  $B$  and such that

$$\rho(\Omega, \langle X(\cdot), Y(\cdot) \rangle \in E, \omega) = b_1 \quad (72)$$

for all  $\omega \in \Omega$ , as this is supposed to be known or easy to verify as an a priori evidence,

$$\rho(\Omega, Y(\cdot) \in B, \omega) = b_2, \quad (73)$$

for all  $\omega \in A_X \cup (\Omega - A^X)$ , as in this case the decision whether  $Y(\omega) \in B$  or not can be converted into the decision whether  $X(\omega) \in S_{X,A}$  or  $X(\omega) \in \mathcal{S} - S^{X,A}$ , finally,

$$\rho(\Omega, Y(\cdot) \in B, \omega) = b_3 \quad (74)$$

for all  $\omega \in A^X - A_X$ , as in this case such a conversion to a decision problem concerning only the observable values of the random variable  $X$  is not possible. Perhaps more intuitively, we could also take  $B = \mathcal{N} \cup \{\infty\}$  with the usual ordering,  $b_1 = 0$ ,  $0 < b_2 < \infty$ ,  $b_3 = \infty$ . Hence,  $A_X$  and  $A^X$  are  $b_2$ -subsets of  $\Omega$  so that  $\langle A_X, A^X \rangle$  is a  $b_2$ -rough set generated, in  $\Omega$ , by the predicate  $P_1(\cdot)$  and reference system  $\rho$ . On the other side,  $A$  itself is a  $b_3$ -subset, but not  $b_2$ -subset of  $\Omega$ , hence, ascribing the status of existence only to  $b_2$ -subsets we could say that  $A$  "does not exist".

When saying, above, that the values of the random variable  $X$  are observable, we have to abandon, for a while and for the sake of simplicity, the basic idea of Dempster-Shafer approach according to which not the value  $X(\omega) \in \mathcal{S}$  itself, but rather the value  $\tilde{S}(\omega) \subset \mathcal{S}$  is observable, our evidence being enriched by the assumption that  $X(\omega) \in \tilde{S}(\omega)$ . Following this way of reasoning and given  $B \subset \mathcal{H}$ , we may define:

$$A_0 = A_0(B) = \left\{ \omega : \omega \in \Omega, \langle X(\omega), Y(\omega) \rangle \in E, X(\omega) \in \tilde{S}(\omega), Y(\omega) \in B \right\}. \quad (75)$$

If  $\tilde{S}(\omega) \subset S_{X,A}$ , then  $X(\omega) \in \tilde{S}(\omega)$  yields that  $X(\omega) \in S_{X,A}$  and we can prove that  $Y(\omega) \in B$  using the same way of reasoning as above. If  $\tilde{S}(\omega) \subset \mathcal{S} - S^{X,A}$ , then  $X(\omega) \in S^{X,A}$  and  $Y(\omega) \in \mathcal{H} - B$  follows. So, we may set

$$A_S = \left\{ \omega : \omega \in \Omega, \tilde{S}(\omega) \subset S_{X,A}, \langle X(\omega), Y(\omega) \rangle \in E, X(\omega) \in \tilde{S}(\omega) \right\}, \quad (76)$$

$$A^S = \left\{ \omega : \omega \in \Omega, \tilde{S}(\omega) \subset \mathcal{S} - S^{X,A}, \langle X(\omega), Y(\omega) \rangle \in E, X(\omega) \in \tilde{S}(\omega) \right\}. \quad (77)$$

Evidently,  $A_S \subset A_0 \subset A^S$ , and if  $X(\omega) \in \tilde{S}(\omega)$  for all  $\omega \in \Omega$  or if we restrict the sets  $A_S$ ,  $A_0$ , and  $A^S$  to the subset of  $\Omega$  satisfying this demand, then also  $A_S \subset A_X$  and  $A^S \subset A^X$ . So,  $\langle A_S, A^S \rangle$  is a rough set with respect to  $A$ , which can be defined as  $b_2$ -rough set using the linearly ordered set  $B$  as above and a monotonous reference system  $\rho_0$  taking its values in  $B$  and such that (72) holds for each  $\omega \in \Omega$  and for the predicate  $\langle (X(\cdot), Y(\cdot)) \in E \rangle \underline{\text{et}} (X(\omega) \in \tilde{S}(\omega))$ , (73) holds for  $\omega \in A_S \cup (\Omega - A^S)$ , as in this case  $Y(\omega) \in B$  can be proved or disproved using the observable facts that  $\tilde{S}(\omega) \subset S_{X,A}$  or  $\tilde{S}(\omega) \subset \Omega - S^{X,A}$ . Finally, (74) holds for  $\omega \in A^S - A_S$ , as in this case we cannot use the observable values of  $\tilde{S}(\omega)$  when deciding about the membership of  $Y(\omega)$  in  $B$ . So,  $A_S$  and  $A^S$  are  $b_2$ -subsets of  $\Omega$ ,  $\langle A_S, A^S \rangle$  is a  $b_2$ -rough set with respect to  $A$ , the predicate  $\left[ \langle (X(\cdot), Y(\cdot)) \in E \rangle \underline{\text{et}} (X(\omega) \in \tilde{S}(\omega)) \underline{\text{et}} (Y(\omega) \in B) \right]$ , and  $\rho$ .  $A$  itself is a

$b_3$ -subset, but not a  $b_2$ -subset with respect to the same predicate and reference system. The values  $Pr\left(\left\{\omega : \tilde{S}(\omega) \subset S_{X,A}\right\}\right)$  and  $1 - Pr\left(\left\{\omega : \tilde{S}(\omega) \subset S - S^{X,A}\right\}\right)$  may serve as a probabilistic numerical quantification of the rough set  $\langle A_S, A^S \rangle$ .

## 7. FROM ROUGH SETS TO FUZZY SETS

As we have already introduced, rough sets have been conceived as a tool for describing and handling uncertainty under a very poorly structured basic knowledge enabling just a three-valued classification. Let us consider, in this chapter, connections between the two interpretations of rough sets investigated above, and the so called fuzzy sets, another well-known non-probabilistic apparatus for indeterminism and uncertainty processing. The reader is supposed to be familiar with the notion and basic properties of fuzzy sets, cf. e. g., Novák [8] as a good source.

The most straightforward way from rough sets to fuzzy sets is as follows. A rough set  $\langle V_*, V^* \rangle$ ,  $V_* \subset V^* \subset X$ , uniquely defines the fuzzy set  $\bar{V}$  such that, denoting by  $\mu(\bar{V})(x)$  the degree of membership of  $x \in X$  in  $\bar{V}$ , we have

$$\begin{aligned}\mu(\bar{V})(x) &= 1, & x \in V_*, \\ \mu(\bar{V})(x) &= \frac{1}{2}, & x \in V^* - V_*, \\ \mu(\bar{V})(x) &= 0, & x \in X - V^*.\end{aligned}\tag{78}$$

However, as proved by Pawlak [14] and quoted by Wygralak [18], this translation is not one-to-one in the sense that some information is lost when transforming  $\langle V_*, V^* \rangle$  into  $\bar{V}$ .

Let  $Y = \langle Y_*, Y^* \rangle$  and  $Z = \langle Z_*, Z^* \rangle$  be rough sets defined by two classical or crisp sets  $Y_0, Z_0$  and by an equivalence relation in  $X$  according to (1), let  $[x]$  denote the corresponding equivalence class for  $x \in X$ . Set, again according to (1),

$$\begin{aligned}(Y_0 \cup Z_0)_* &= \{x \in X, [x] \subset Y_0 \cup Z_0\}, \\ (Y_0 \cup Z_0)^* &= \{x \in X, [x] \cap (Y_0 \cup Z_0) \neq \emptyset\}, \\ (Y_0 \cap Z_0)_* &= \{x \in X, [x] \subset Y_0 \cap Z_0\}, \\ (Y_0 \cap Z_0)^* &= \{x \in X, [x] \cap (Y_0 \cap Z_0) \neq \emptyset\}.\end{aligned}\tag{79}$$

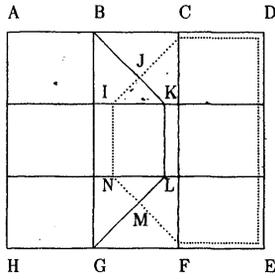
Two binary operations  $\sqcup$  and  $\sqcap$  on rough sets can be defined in this way:

$$\begin{aligned}Y \sqcup Z &= \langle (Y_0 \cup Z_0)_*, (Y_0 \cup Z_0)^* \rangle, \\ Y \sqcap Z &= \langle (Y_0 \cap Z_0)_*, (Y_0 \cap Z_0)^* \rangle.\end{aligned}\tag{80}$$

Pawlak proved [14] that considering the corresponding fuzzy sets,

$$\begin{aligned}\mu(\overline{Y \sqcup Z})(x) &\neq \max\{\mu(\bar{Y})(x), \mu(\bar{Z})(x)\}, \\ \mu(\overline{Y \sqcap Z})(x) &\neq \min\{\mu(\bar{Y})(x), \mu(\bar{Z})(x)\}\end{aligned}\tag{81}$$

in general, i.e. for all  $x \in X$ . The equality in (81) holds for all  $x \in X$  iff  $(Y_0 \cup Z_0)_* = Y_{0*} \cup Z_{0*}$  and  $(Y_0 \cap Z_0)^* = Y_0^* \cap Z_0^*$ , which is not the case in general, only  $(Y_0 \cup Z_0)^* = Y_0^* \cup Z_0^*$  and  $(Y_0 \cap Z_0)_* = Y_{0*} \cap Z_{0*}$  hold in all cases. The following figure illustrates the situation more intuitively, cf. [18].



Let  $X = ADEH$  be the universum, let  $Y_0 = ABKLGH$ ,  $Z_0 = CDEFNI$ , let the equivalence relation in  $X$  be obtained by the nine squares, i.e., points in each square are inseparable from each other. Hence,  $\langle Y_{0*}, Y_0^* \rangle = \langle ABGH, ACFH \rangle$ ,  $\langle Z_{0*}, Z_0^* \rangle = \langle CDEF, BDEG \rangle$ . Let  $x \in IKLN$ , let  $y \in IJK \cup LMN$ , then

$$\mu(\overline{Y})(x) = \mu(\overline{Y})(y) = \mu(\overline{Z})(x) = \mu(\overline{Z})(y) = \frac{1}{2}, \quad (82)$$

but

$$1 = \mu(\overline{Y \sqcup Z})(x) \neq \mu(\overline{Y \sqcup Z})(y) = \frac{1}{2}. \quad (83)$$

So, it is not possible to define a binary numerical operation  $\cup_0$ , mapping  $\{0, \frac{1}{2}, 1\} \times \{0, \frac{1}{2}, 1\}$  into  $\{0, \frac{1}{2}, 1\}$  in such a way that, for all  $z \in X$

$$\mu(\overline{Y \sqcup Z})(z) = (\mu(\overline{Y})(z)) \cup_0 (\mu(\overline{Z})(z)). \quad (84)$$

Or, supposing this were possible, we obtain, for  $x$  and  $y$  as above,

$$(\mu(\overline{Y})(x)) \cup_0 (\mu(\overline{Z})(x)) = 1, \quad \text{hence} \quad \left(\frac{1}{2}\right) \cup_0 \left(\frac{1}{2}\right) = 1, \quad (85)$$

but, at the same time,

$$(\mu(\overline{Y})(y)) \cup_0 (\mu(\overline{Z})(y)) = \frac{1}{2}, \quad \text{hence} \quad \left(\frac{1}{2}\right) \cup_0 \left(\frac{1}{2}\right) = \frac{1}{2}, \quad (86)$$

so that we have arrived at a contradiction. An analogous straightforward consideration easily proves the impossibility to define the dual binary operation  $\cap_0$ . In other words

not extensional) with respect to three-valued fuzzy sets in the sense that the values of membership functions for  $\overline{Y \sqcup Z}$  and  $\overline{Y \sqcap Z}$  cannot be defined as functions of membership values for  $\overline{Y}$  and  $\overline{Z}$ .

We do not aim to go into more details as far as this simple relation between rough sets and fuzzy sets is concerned, turning our attention to another bridge connecting the two notions and opened due to the alternative approach to rough sets as explained above.

Recall that reference systems are mappings taking their values in a partially ordered set  $\langle B, \preceq \rangle$ , and denote by  $\mathcal{P}(B)$  the power set consisting of all subsets of  $B$ . A natural partial ordering on  $\mathcal{P}(B)$  is that generated by the set-theoretical inclusion. Set, for each  $b \in B$ ,

$$S_1(b) = \{c : c \in B, b \preceq c\}, \quad (87)$$

$$S_2(b) = \{c : c \in B, \underline{\text{non}}(c \preceq b)\}. \quad (88)$$

Evidently,  $b \in S_1(b)$  for each  $b \in B$ , if  $I$  is the maximal element of  $\langle B, \preceq \rangle$ , i. e.,  $b \preceq I$  for each  $b \in B$ , then  $S_2(I) = \emptyset$  (the empty subset of  $B$ ). Now, given  $A \subset X$ ,  $P \in \underline{\text{Pred}}$ , and  $x \in X$ , two fuzzy subsets  $\overline{A|P^1}$ ,  $\overline{A|P^2}$  of  $A$  may be defined with membership function  $\mu$  taking its values in  $\mathcal{P}(B)$ . Namely, for  $i = 1, 2$ ,

$$\mu(\overline{A|P^i})(x) = S_i(\rho(A, P, x)), \quad (89)$$

if  $x \in A$  and  $\rho(A, P, x)$  is defined,

$$\mu(\overline{A|P^i})(x) = \emptyset \quad (90)$$

otherwise, i. e. if either  $x \in X - A$  or if  $\rho(A, P, x)$  is not defined.

Both the membership functions  $\mu(\overline{A|P^i})$ ,  $i = 1, 2$ , can be easily proved to be monotonous in an intuitive sense.

**Fact 4.** If  $\rho(A, P, x_1) \preceq \rho(A, P, x_2)$ , or if  $\rho(A, P, x_2)$  is not defined, then  $\mu(\overline{A|P^i})(x_1) \supseteq \mu(\overline{A|P^i})(x_2)$  for both  $i = 1, 2$ .

*Proof.* If  $\rho(A, P, x_2)$  is not defined, then  $\mu(\overline{A|P^i})(x) = \emptyset$  and the assertion is trivial, let  $\rho(A, P, x_1) = b_1 \preceq b_2 = \rho(A, P, x_2)$  be both defined. If  $b_2 \preceq c$ , then  $b_1 \preceq c$  as well due to the transitivity of partial ordering relations, so that  $S_1(b_1) \supseteq S_1(b_2)$ . On the other hand side,  $c \preceq b_1$  implies  $c \preceq b_2$ , so that  $\underline{\text{non}}(c \preceq b_2)$  implies  $\underline{\text{non}}(c \preceq b_1)$ ; hence,  $S_2(b_1) \supseteq S_2(b_2)$  as well. The assertion is proved.  $\square$

**Fact 5.** Let  $X, A \subset X$ ,  $\underline{\text{Pred}}$  is as above, let  $\rho$  be a monotonous reference system, let  $P_1, P_2 \in \underline{\text{Pred}}$ , let  $A_i = A|P_i$ ,  $i = 1, 2$ , let  $\rho(A, P_1, x)$ ,  $\rho(A, P_2, x)$  be defined, then

$$\begin{aligned} \mu(\overline{A_1 \cup A_2^1})(x) &\supseteq \sup \left\{ \mu(\overline{A_1^1})(x), \mu(\overline{A_2^1})(x) \right\} = \\ &= \mu(\overline{A_1^1})(x) \cup \mu(\overline{A_2^1})(x), \end{aligned} \quad (91)$$

$$\begin{aligned} \mu(\overline{A_1 \cup A_2^2})(x) &= \sup \{ \mu(\overline{A_1^2})(x), \mu(\overline{A_2^2})(x) \} = \\ &= \mu(\overline{A_1^2})(x) \cup \mu(\overline{A_2^2})(x), \end{aligned} \quad (92)$$

$$\begin{aligned} \mu(\overline{A_1 \cap A_2^1})(x) &= \inf \{ \mu(\overline{A_1^1})(x), \mu(\overline{A_2^1})(x) \} = \\ &= \mu(\overline{A_1^1})(x) \cap \mu(\overline{A_2^1})(x), \end{aligned} \quad (93)$$

$$\begin{aligned} \mu(\overline{A_1 \cup A_2^1})(x) &\subseteq \inf \{ \mu(\overline{A_1^2})(x), \mu(\overline{A_2^2})(x) \} = \\ &= \mu(\overline{A_1^2})(x) \cap \mu(\overline{A_2^2})(x). \end{aligned} \quad (94)$$

Proof. For both  $j = 1, 2$ ,

$$\begin{aligned} \mu(\overline{A_1 \cup A_2^j})(x) &= \mu(\overline{A_1 \mid P_1 \cup A \mid P_2^j})(x) = \mu(\overline{A \mid (P_1 \vee \text{el} P_2^j)})(x) = \\ &= S_j(\rho(A, P_1 \vee \text{el} P_2, x)) = S_j(\rho(A, P_1, x) \wedge \rho(A, P_2, x)) \end{aligned} \quad (95)$$

due to (9). But, in general, for each  $b_1, b_2 \in B$ ,

$$\begin{aligned} S_1(b_1 \wedge b_2) &= \{c : c \in B, b_1 \wedge b_2 \preceq c\} \supseteq \\ &\supseteq \{c : c \in B, b_1 \preceq c\} \cup \{c : c \in B, b_2 \preceq c\} = S_1(b) \cup S_1(b), \end{aligned} \quad (96)$$

as  $b_1 \preceq c$  or  $b_2 \preceq c$  imply  $b_1 \wedge b_2 \preceq c$ . For  $S_2$  we obtain

$$S_2(b_1 \wedge b_2) = \{c : c \in B, \underline{\text{non}}(c \preceq b_1 \wedge b_2)\}. \quad (97)$$

But, due to the properties of  $\wedge$  operation, cf. (6) and (7),

$$\{c : c \in B, c \preceq b_1\} \cap \{c : c \in B, c \preceq b_2\} = \{c : c \in B, c \preceq b_1 \wedge b_2\}, \quad (98)$$

so that, by de Morgan rules,

$$\begin{aligned} \{c : c \in B, \underline{\text{non}}(c \preceq b_1 \wedge b_2)\} &= \\ &= \{c : c \in B, \underline{\text{non}}(c \preceq b_1)\} \cup \{c : c \in B, \underline{\text{non}}(c \preceq b_2)\}, \end{aligned} \quad (99)$$

hence,

$$S_2(b_1 \wedge b_2) = S_2(b_1) \cup S_2(b_2). \quad (100)$$

So,

$$\begin{aligned} S_1(\rho(A, P_1, x) \wedge \rho(A, P_2, x)) &\supseteq S_1(\rho(A, P_1, x)) \cup S_1(\rho(A, P_2, x)) = \\ &= \mu(\overline{A_1^1})(x) \cup \mu(\overline{A_2^1})(x), \end{aligned} \quad (101)$$

and

$$\begin{aligned} S_2(\rho(A, P_1, x) \wedge \rho(A, P_2, x)) &= S_2(\rho(A, P_1, x)) \cup S_2(\rho(A, P_2, x)) = \\ &= \mu(\overline{A_1^2})(x) \cup \mu(\overline{A_2^2})(x). \end{aligned} \quad (102)$$

Similarly, for both  $j = 1, 2$ ,

$$\begin{aligned} \mu(\overline{A_1 \cup A_2^j})(x) &= \mu(\overline{A | P_1 \cap A | P_2^j})(x) = \mu(\overline{A | (P_1 \text{ et } P_2^j)})(x) = (103) \\ &= S_j(\rho(A, P_1 \text{ et } P_2, x)) = S_j(\rho(A, P_1, x) \vee \rho(A, P_2, x)) \end{aligned}$$

due to (8). But, in general, for each  $b_1, b_2 \in B$ ,

$$\begin{aligned} S_1(b_1 \vee b_2) &= \{c : c \in B, b_1 \vee b_2 \leq c\} = (104) \\ &= \{c : c \in B, b_1 \leq c\} \cap \{c : c \in B, b_2 \leq c\} = S_1(b_1) \cap S_2(b_2) \end{aligned}$$

due to (7). For  $S_2$  we obtain

$$S_2(b_1 \vee b_2) = \{c : c \in B, \underline{\text{non}}(c \leq b_1 \vee b_2)\}. \quad (105)$$

But, due to (6) and (7),

$$\{c : c \in B, c \leq b_1\} \cup \{c : c \in B, c \leq b_2\} \subset \{c : c \in B, c \leq b_1 \vee b_2\}, \quad (106)$$

so that

$$\begin{aligned} \{c : c \in B, \underline{\text{non}}(c \leq b_1 \vee b_2)\} &\supset (107) \\ &\supset \{c : c \in B, \underline{\text{non}}(c \leq b_1)\} \cup \{c : c \in B, \underline{\text{non}}(c \leq b_2)\}, \end{aligned}$$

hence,

$$S_2(b_1 \vee b_2) \subseteq S_2(b_1) \cap S_2(b_2). \quad (108)$$

So,

$$\begin{aligned} S_1(\rho(A, P_1, x) \vee \rho(A, P_2, x)) &= S_1(\rho(A, P_1, x)) \cap S_1(\rho(A, P_2, x)) = (109) \\ &= \mu(\overline{A_1^1})(x) \cap \mu(\overline{A_2^1})(x), \end{aligned}$$

and

$$\begin{aligned} S_2(\rho(A, P_1, x) \vee \rho(A, P_2, x)) &\subseteq S_2(\rho(A, P_1, x)) \cap S_2(\rho(A, P_2, x)) = (110) \\ &= \mu(\overline{A_1^2})(x) \cap \mu(\overline{A_2^2})(x). \end{aligned}$$

Hence, (16), (17) and, consequently, whole the assertion are proved.  $\square$

(Received September 14, 1990.)

#### REFERENCES

- [1] M. Davis: Computability and Unsolvability. McGraw-Hill, New York 1958.
- [2] A. A. Fraenkel and Y. Bar-Hillel: Foundations of Set Theory. North-Holland, Amsterdam 1959 (Russian translation: Mir, Moscow 1966).
- [3] P. R. Halmos: Measure Theory. D. van Nostrand, New York 1950 (Russian translation: Izd. Innostr. Lit., Moscow 1953).

- [4] M. Kubát: More attention to the rough set theory. In: Aplikace umělé inteligence AI 89, ÚISK Prague, 1989, pp. 361 – 368.
- [5] T. S. Kuhn: The Structure of Scientific Revolutions. Oxford Univ. Press, Oxford 1980.
- [6] K. Kuratowski and A. Mostowski: Teoria mnogości (Set Theory). Second edition. PWN, Warsaw 1966.
- [7] P. Maddy: Believing the axioms. J. Symbolic Logic 53 (1988), No. 2, pp. 481 – 511, No. 3, pp. 736 – 764.
- [8] V. Novák: Fuzzy Sets and their Applications. Adam Hilger, Bristol – Philadelphia, SNTL, Prague 1989.
- [9] E. Orłowska: Semantics of knowledge operators. Bull. Polish Acad. Sci. Math. 35 (1987), 5-6, 255 – 263.
- [10] E. Orłowska: Algebraic approach to database constraints. Fund. Inform. 10 (1987), 1, 57 – 68.
- [11] E. Orłowska: Representation of vague information. Inform. Systems 13 (1988), 2, 167 – 174.
- [12] Z. Pawlak: Information systems – theoretical foundations. Inform. Systems 6 (1981), 2, 205 – 218.
- [13] Z. Pawlak: Rough sets. Internat. J. Comput. Inform. Sci. 11 (1982), 3, 341 – 356.
- [14] Z. Pawlak: Rough Sets and Fuzzy Sets. Institute of Computer Science, Polish Academy of Sciences, Report No. 540, Warsaw 1984.
- [15] J. Pearl: Probabilistic Reasoning in Intelligent Systems – Networks of Plausible Inference. Morgan Kaufmann Publishers, Inc., San Mateo, Calif. 1988.
- [16] G. Shafer: A Mathematical Theory of Evidence. Princeton Univ. Press, Princeton 1976.
- [17] G. Shafer: Probability judgment in artificial intelligence and expert systems. Statist. Sci. 2 (1987), 1, 3 – 4.
- [18] M. Wygralak: Some remarks on rough and fuzzy sets. Busefal 21 (1985), 1, 42 – 48.

*RNDr. Ivan Kramosil, DrSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation – Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8, Czechoslovakia.*