UDC: 007.52:519.21:517.518.2(047)=20
COSATI: 09-02, 12-01, 14-02

Uncertainty phenomena and formalisms of their modeling in expert systems

Dragan Šaletić, MSc (Eng)[1)]
Dušan Velašević, PhD (Eng)[2)]

Different types of uncertainty and their relations are pointed out in the paper. The short review of mathematical formalisms used for modeling uncertainty in expert systems is given. The difference between probability and fuzziness, expressed also in a formal mathematical manner, is shown. Conditions specifying applications of considered approaches to uncertainty modeling, are defined, thus making easier the modeling of uncertainty in expert and other systems of artificial intelligence. The possibilities of further research work in the domain are considered.

Key words: **computer science, artificial intelligence, expert systems, uncertainty, fuzzy logic, probability, theory of rough sets.**

List of notation and symbols

$S=\{s_1, s_2, \dots, s_n\}$	– set, domain of discourse, of elements $s_1, s_2, \dots, s_n$
s	– generic element in the set S
$\mathcal{P}(S)$	– power set of the set S, the set of ordinary subsets of the set S
V	– variable that can take either crisp values or fuzzy values
μ	– measure of belief
$\mu_A(x)$	– membership function of the element x with the respect to the fuzzy set A
χ_B	– characteristic function of the crisp set B
$f : C \to D$	– function f that maps elements of the set C onto the set D
$\wedge$	– conjunction, and
$\vee$	– disjunction, or
$\Rightarrow$	– implication
$\{s \mid F\}$	– set of elements with the feature F
P	– probability
Π	– possibility
π	– possibility distribution
$\mathcal{N}$	– necessity
$\sup A$	– supremum of A , the least upper bound of A
$\inf A$	– infimum of A , the greatest lower bound of A
m	– basic probability assignment
Cr	– credibility function
Pl	– plausibility function
$lower(\Sigma)$	– lower approximation (also known as the positive region, $pos(\Sigma)$) of the set Σ
$[s]_R$	– equivalence class of the relation R
$upper(\Sigma)$	– upper approximation of the set Σ
$bnd(\Sigma)$	– boundary region of the set Σ
$neg(\Sigma)$	– negative region of the Σ

Introduction

AS applicability of some artificial intelligence systems (i.e. expert systems, neural networks, etc.), in real problem solving increases, it is more obvious that the knowledge needed for finding solutions of these problems is inherently uncertain. Many practical problems are pervaded by uncertainty [1]. Almost all information is subject to uncertainty. Uncertainty may arise from inaccurate or incomplete information, from linguistic imprecision, from disagreement between information sources, or from an insufficiently defined or ill-defined problem. Abilities to process uncertain information and to reason on the basis of insufficient knowledge are determining features of intelligent behavior in an *uncertain*, i.e. complex and dynamical, environment. Those are the reasons that make uncertainty modeling the important research field in the domain of artificial intelligence in the last several decades and nowadays. Uncertainty modeling problems become important with arising of advanced information systems, equipped with some possibilities of reasoning. Hence this field is the topic of many research projects, conferences and papers. This paper points out to different types of uncertainty, and to relations between them. The short review of mathematical formalisms used for modeling uncertain information in rule-based expert systems [2],[3],[4] in uncertain environments, is given, with the aim to clarify possible disagreements about basic assumptions and appropriate applicabilities of uncertainty models in expert systems. Disagreements exist in foreign [5], as well as in domestic scientific circles.

1) Military Technical Institute (VTI) of the Yugoslav Army, Katanićeva 15, 11000 Beograd
2) Faculty of Electrical Engineering, Bulevar kralja Aleksandra 73, 11000 Beograd

The difference between probability and fuzziness, expressed also in a formal mathematical manner, is shown. Conditions specifying applications of considered approaches to uncertainty modeling, are defined, and modeling of uncertainty in expert and other systems of artificial intelligence is thus made easier. The possibilities of further research work in the domain are considered.

The following uncertainty phenomena and methods of their modeling in expert systems in uncertain environments are considered:

- randomness,
- fuzziness, and
- roughness.

These are three different cognitive processes, which exist each *per se*.

Randomness is modeled in expert systems by using:

- probability theory [6], [3];
- heuristic approximation of probability, so called certainty factors [3];
- the Dempster–Shafer theory of evidence [6], [16] - [22].

Fuzziness is modeled by using:

- the theory of fuzzy sets (i.e. fuzzy logic) [7], [8], [9], [10], [11], [6];
- the theory of possibility [12].

Roughness is modeled by the theory of rough sets [23], [24], [25], [26].

In the sequel, the definitions, as well as the survey of the basic features of considered uncertainty types and their models, are given. Some other modeling formalisms, such as nonmonotonic logics and default reasoning, and qualitative versions of probability, such as the Spohn calculus and kappa-calculus, have not been considered here because of existing controversies about those approaches. It is pointed out to kinds of uncertainties which some uncertainty types represent, as well as to relations between different uncertainty types, which makes easier the applicability of uncertainty modeling in expert and other artificial intelligence systems.

The starting point for uncertainty consideration can be a measure of belief [6]. Formally, a belief measure is defined on nonempty set $S = \{s_1, s_2, ..., s_n\}$, $n \in N$, N the set of natural numbers. Let V be a variable which can take values in the set S. When an uncertainty exists, i.e. in situations when the exact value of the variable V is not known, the most that can be done is to try to formulate the knowledge about the variable V, using a convenient mathematical form. One such form is a measure of belief. Its convenience is in the fact that it can be used in representations of different types of uncertainty. Assume $\{X_n \mid n \in N\}$ is a sequence of sets $X_1, X_2, ..., X_n$, ($n = 1,2, ...$, i.e. $n \in N$).

A field of subsets of a set $S \neq \varnothing$ is a system $\mathcal{F}$ of the subsets of S, containing S and $\varnothing$, and closed* under the set operations union $\cup$ and complement $\neg$. (Due to commutativity, associativity and distributivity of the set operations and due to the existence of neutral and unity elements, the set $\mathcal{F}$, equipped with the union $\cup$ and the complement $\neg$, is a Boolean algebra with respect to (these) set-theoretic operations.). The system $\mathcal{F}$ is a *σ-field* if for each sequence $\{X_n \mid n \in N\}$ of the elements of $\mathcal{F}$, the union of these elements, $\cup_{n\in N} X_n$, is an element of $\mathcal{F}$. If to each set $A \in \mathcal{F}$ corresponds a finite real number or $+\infty$, i.e., if Φ is a set function which maps $\mathcal{F}$ onto a real number or $+\infty$, it is said that on $\mathcal{F}$ a set function $\Phi(A)$ is defined

$$\Phi : \mathcal{F} \rightarrow \text{ real number or } +\infty$$

Definition 1. *A belief measure* on $(S, \mathcal{F})$ is a set function $\mu : \mathcal{F} \rightarrow [0,1]$, such that

$$\mu(\varnothing) = 0, \qquad \mu(S) = 1 \tag{1}$$

$$X, Y \in \mathcal{F} \text{ and } X \subseteq Y \text{ implies } \mu(X) \leq \mu(Y) \quad \text{(monotonicity)} \tag{2}$$

The belief measure sometimes is also called *a generalized fuzzy measure* [13], [10], [27].

Within the framework of using the belief measure to represent information about an uncertain variable V, μ_E can be interpreted as a measure associated with our belief that the value of V is contained in the subset $E \subset S$, i.e. as a confidence that $V \in E$.

Probabilistic model of uncertainty

Probability is a kind of a quantitative representation of one uncertainty type – randomness. Probability represents a degree of belief that a proposition is a true, or that an event is going to happen, and that degree is given by a number the value of which is between 0 and 1. In the frequentist view, probability of an event is the frequency of the event occurring in a large number of similar trials.

A probability can be considered from the point of view of the measure theory, as well [6].

Definition 2. A belief measure μ on $(S, \mathcal{F})$ is a *finitely additive probability* if $X, Y \in \mathcal{F}$ and $X \cap Y = \varnothing$ implies $\mu(X \cup Y) = \mu(X) + \mu(Y)$. A measure μ is a *σ - additive probability* (or simply *probability*) if its domain $\mathcal{F}$ is a σ-field, and for each sequence $\{X_n \mid n \in N\}$ of pairwise disjoint elements from $\mathcal{F}$ (i.e. $X_i \cap X_j = \varnothing$ for $i \neq j$), it holds

$$\mu(\cup_{n \in N} X_n) = \Sigma_{n \in N}\ \mu(X_n) \tag{3}$$

The basic properties of probability are [6]:

1. Let μ be a finitely additive probability on $(S, \mathcal{F})$. Then for each finite sequence $X_1, X_2, .., X_n$ of pairwise disjoint elements of $\mathcal{F}$ it holds

$$\mu\left(\bigcup_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mu(X_i) \tag{4}$$

2. Let μ be a *probability* on $(S, \mathcal{F})$ and let $\{X_n \mid n \in N\}$ be a sequence of elements of $\mathcal{F}$. If the sequence is increasing (i.e. $X_i \subseteq X_{i+1}$ for each i), then $\mu(\cup_{i \in N} X_i) = \sup_{i \in N} \mu(X_i)$. If the sequence is decreasing (i.e. $X_i \supseteq X_{i+1}$ for each i), then $\mu(\cap_{i \in N} X_i) = \inf_{i \in N} \mu(X_i)$.

Let S be finite and let $\mathcal{F}$ be the set $\mathcal{P}(S)$, the set of all subsets of S, the power set of S.

3. Each finitely additive probability on $(S, \mathcal{F})$ is σ- additive.
4. Each probability μ on $(S, \mathcal{F})$ is uniquely given by its value of singletons*, i.e.: if $\mu_0: S \rightarrow [0,1]$ is such that $\Sigma_{s \in S}\ \mu_0(s) = 1$, then a

*) A set S is *closed under an operation* * if a result of the operation* applied on any two elements from S is again an element of S.

*) A *singleton* is an elementary event, a set of one element.

unique probability μ exists on $(S, \mathcal{F})$, such that $\mu(\{s\}) = \mu_0(s)$ for each $s \in S$.

Therefore, a probability is an *additive measure*. In the probability theory, the additive measure μ is usually denoted as P. The basic set S consists of elementary events or propositions, and $\mathcal{F}$ is usually taken as the power set of S, denoted as $\mathcal{P}$, which has a *ring* structure (in the sense of set theory), i.e. it holds ($\neg$ is negation)

$$\text{if } p \in \mathcal{P}, \text{ then } \neg p \in \mathcal{P} \tag{5}$$

$$\text{if } p \in \mathcal{P} \text{ and } q \in \mathcal{P}, \text{ then } p \wedge q \in \mathcal{P} \tag{6}$$

and then the basic statements can be formulated that define a probability in an alternative approach to the approach represented by the expressions (1–4).

Expressions (5) and (6) state that the ring $\mathcal{P}$ is closed under operations of negation and conjunction (it is easy to prove that the same holds for disjunction and difference).

Definition 3. Suppose that the set $\mathcal{P}$ is the power set of a basic set S of elementary propositions, which is called the *basic space* (*set*). On the set $\mathcal{P}$ a *probability P* can be defined as a measure, i.e. a function $P: \mathcal{P} \to [0,1]$, such that

1) $P(\varnothing) = 0$, where $\varnothing$ is a proposition which is always false (7)

2) $P(S) = 1$, where S is a proposition which is always true (8)

3) $(\forall p \in \mathcal{P}, \forall q \in \mathcal{P}, p \wedge q = 0) \Rightarrow P(p \vee q) = P(p) + P(q).$ (9)

Expressions (7) and (8) are expressions (1) in the case of a probability.

When it holds $p \wedge q = 0$, the propositions p and q are ″mutually disjoint″, since one of them is false if the other one is true. Based on axioms (7–9) the following consequences hold:

$$\forall p \in \mathcal{P}, \; P(p) + P(\neg p) = 1 \tag{10}$$

(the probabilistic law of the excluded middle)

and the expression (2) in the probabilistic case

$$(p \Rightarrow q) \Rightarrow P(p) \le P(q) \tag{11}$$

Implicitly, an order relation between the propositions is introduced, which corresponds to the inclusion in the set interpretation.

The main form of a probabilistic reasoning in expert systems consists of using the Bayes theorem. On the basis of Bayes's theorem, *a priori* probability of the proposition, and probabilities of new facts, the *a posteriori* probability is inferred (Bayesian reasoning)

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)} \tag{12}$$

where

$P(H|E)$– is the probability that H is true given the evidence E;

$P(H)$– is the probability that H is true;

$P(E|H)$– is the probability of observing the evidence E when H is true;

$P(E)$– is the probability of E.

Bayes's theorem is used in rule-based expert systems, in which there is a rule of the following form

$$\text{IF } E \text{ THEN } H \tag{13}$$

Using expression (12), the probability of the hypothesis H for the given evidence E, can be determined [3].

In such a probabilistic calculus a proposition p and its negation, complement, $\neg p$, are related by expression (10). This is intuitively acceptable in some cases. However, difficulties arise when there is no any *a priori* information about the truth of the proposition p. It looks natural to take $P(p) = P(\neg p) = 0.5$: when there is no knowledge about the probability distribution, the symmetric truth values are given (0.5), i.e. a so-called *principle of insufficient reasoning* [19], is applied. But, assume there are three propositions, p_1, p_2, and p_3, with no knowledge of their validity. Analogously to the case of two mutually disjoint propositions, it seems reasonable to take $P(p_1)=P(p_2)=P(p_3)=0.33...$ But, if p and q from the previous case of two propositions, are now taken as $p = p_1$, $q = p_2 \vee p_3$, it holds that $P(p_1) = 0.5$, which is not in agreement with the symmetric distribution (0.33..), and that results in a contradiction.

The probabilistic calculus demands existence of an *a priori* information. Conditional independence of data must be assumed. Additivity holds – the sum of probabilities that support a hypothesis and are against it, for a given fact, have to equal one. Then the probability theory offers advantages of a well-founded and statistically correct method of inexact reasoning. An interpretation of probabilities as relative frequencies demands a massive population of data as a foundation of relevant statistics. The Bayes model is a good choice for the situation when there is a lot of information, but is too restrictive for the most of real situations. It is often a case in real applications that the needed conditions are not fulfilled, i.e. there is no previous data on probability distribution. (The probabilistic approach is used, for example, in the geological expert system PROSPECTOR [3], [28]). Due to those reasons, the other mathematical models of uncertainty are introduced, among which some are generalizations, in a way, of a concept of probability.

Certainty factors

Certainty factors are a heuristic approximation of the probability.The certainty factor is added to an IF-THEN rule and it expresses the value of belief in the uncertain rule

$$\mathbf{R}_i\text{: IF } E \text{ THEN } H \; (\mathrm{CF}_i) \tag{14}$$

where $\mathbf{R}_i$ denotes the i-th rule in a rule base of n rules, and CF_i denotes a certainty factor of the considered rule. The value of belief of H given E is true, $\mathrm{CF}_{H,E}$ is

$$\mathrm{CF}_{H,E} = \mathrm{CF}_i$$

The rules of the form given by (14) are called single premise rules.

The calculus of certainty factors is developed [3].

If the available evidence E contained in the rule's premise in the single premise rule is uncertain, and if that uncertainty is described by the CF value of the premise, CF_E, the rule has the following form

$$\mathbf{R}_i\text{: IF } E \; (\mathrm{CF}_E) \text{ THEN } H \; (\mathrm{CF}_i) \tag{15}$$

In this case, certainty factor propagation is present. Certainty factor propagation is concerned with establishing the value of belief in the rule's conclusion, $\mathrm{CF}_{H,E}$, when the uncertainty of evidence in the rule's premise exists. For rule

(15), the value of belief in rule's conclusion is given by the following expression

$$CF_{H,E} = CF_E \cdot CF_i \tag{16}$$

where "·" is the multiplication symbol.

The certainty factor propagation technique is used also when the rule base with inference chains is present. The characteristic of an inference chain is that the conclusion of some previous rule supports the premise of a current rule. A rule base which consists of two rules with an inference chain is given by the following

$\mathbf{R}_1$: IF A THEN B (CF_1)

$\mathbf{R}_2$: IF B THEN C (CF_2)

The certainty theory model propagates the certainty value through an inference chain in rules $\mathbf{R}_1$ and $\mathbf{R}_2$ as independent probabilities

$$CF_{C,A} = CF_{C,B} \cdot CF_{B,A} \tag{17}$$

although in the probability theory, in general, $P(C|A) \neq \neq P(C|B) \cdot P(B|A)$.

Example 1. Suppose the following rule base is given

$\mathbf{R}_1$: IF A ($CF_A = 0.7$) THEN B ($CF_1 = 0.8$)

$\mathbf{R}_2$: IF B THEN C ($CF_2 = 0.9$)

The propagation of $CF_A = 0.7$ through the rule $\mathbf{R}_1$ gives the value of belief of the conclusion B for the given A, according to (16), $CF_{B,A} = CF_A\ CF_1 = 0.7 \cdot 0.8 = 0.56$. The further certainty propagation, now through the rule $\mathbf{R}_2$, gives first $CF_B = CF_{B,A} = 0.56$, and then $CF_C = CF_B \cdot CF_2 = =0.56 \cdot 0.9 = 0.504$, or (according to (17)) $CF_C = CF_{C,A} = =CF_{C,B} \cdot CF_{B,A} = 0.9 \cdot 0.56 = 0.504$. The value of belief in conclusion C is 0.504.

In the certainty factor theory a rule with more than one premise in IF part is interpreted either as a rule with conjunctive premises

$$\mathbf{R}_i\text{: IF } E_1\ (CF_{E1}) \text{ AND } E_2\ (CF_{E2}) \text{ AND} \ldots \ldots \text{ AND } E_n\ (CF_{En}) \text{ THEN } H\ (CF_i) \tag{18}$$

or as a rule with disjunctive premises

$$\mathbf{R}_i\text{: IF } E_1\ (CF_{E1}) \text{ OR } E_2\ (CF_{E2}) \text{ OR} \ldots \ldots \text{ OR } E_n\ (CF_{En}) \text{ THEN } H\ (CF_i) \tag{19}$$

In forming a belief in a hypothesis H, supported by conjunctive or disjunctive rules, in the certainty factor theory the conditional independence of evidence is assumed. The certainty factor CF_E of the IF part of the rule is determined by taking the minimum uncertainty value over the certainty factors of the premises in IF part, (CF_{Ei}), $i = 1,\ldots, n$, in the case of a rule with conjunctive premises

$$CF_E = \min[CF_{E1}, CF_{E2}, \ldots, CF_{En}] \tag{20}$$

and then the value of belief in the rule's conclusion is determined by expression (16).

The certainty factor CF_E of the IF part of the rule is determined by taking the maximum uncertainty value over the certainty factors of the premises in IF part, (CF_{Ei}), $i = 1,\ldots, n$, in the case of a rule with disjunctive premises

$$CF_E = \max[CF_{E1}, CF_{E2}, \ldots, CF_{En}] \tag{21}$$

This policy adopted in the certainty factor theory is in accordance with the canons of the fuzzy set theory [7], [28].

Example 2. Suppose the following rule is given

$\mathbf{R}_1$: IF A ($CF_A = 0.9$) AND

B (CF

Definition 4. Assume S is a set, a universe of discourse, of objects generically denoted as s. *A fuzzy set A* in S, $A \subset S$, is the set of ordered pairs

$$A = \{ (s, \mu_A(s)) \mid s \in S \} \tag{22}$$

specified by a grade of membership μ_A, which is used to associate with each element (each point) $s \in S$, its grade of membership to the set A, $\mu_A(s)$. A function $\mu_A(s)$ is called *a membership function* of the fuzzy set A, where $\mu_A(s)$ takes values in the unit interval [0,1].

In a more general case, $\mu_A(s)$ can take values in a partially ordered set [7]. A discrete membership function is depicted in Fig.1, and a continuous membership function is depicted in Fig.2. In the case of traditional, crisp sets, the concept of membership function is equivalent to the well--known concept of a characteristic function (usually denoted by χ : $\mu_A(s) = \chi_A(s)$), which takes values 1, when s belongs to A, or 0, when s does not belongs to A.

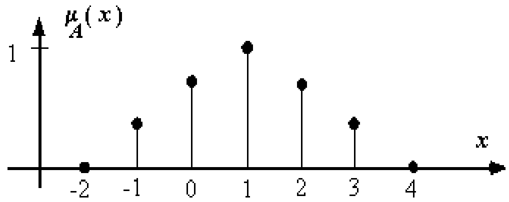

Figure 1. A discrete membership function of the fuzzy set " approximately 1"

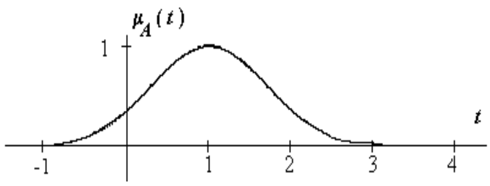

Figure 2. A continuous membership function of the fuzzy set " close to 1"

The most important modeling tool, based on the theory of fuzzy sets, is a fuzzy expert system. The base elements of that system are fuzzy expert rules and approximate (fuzzy) reasoning.

A fuzzy IF – THEN rule (fuzzy implication, fuzzy rule, fuzzy conditional statement) is an expert IF – THEN rule of the following form

$$\text{IF } s \text{ is } A, \text{ THEN } y \text{ is } B \tag{23}$$

where s and y are variables taking values in universes of discourse S and Y, respectively, and A and B are fuzzy sets in S and Y. The interpretation of fuzzy rule enables calculation of conclusions from the sequence of such rules which makes a rule base, that is, makes fuzzy reasoning possible. Fuzzy (approximate) reasoning is a technique of automated reasoning, which gives the procedure of getting an inference from fuzzy IF-THEN rules and known facts.

An approach based on fuzzy logic has been used in the first fuzzy expert systems, implemented by Mamdani's group at Queen Marry College in London. Fuzzy logic controllers have been developed from those expert systems.

The theory of fuzzy sets provides a mathematical framework for encompassing gradualness in computer implementations of reasoning. This gradualness should not necessarily be numeric one, it can be founded on order, on the mathematical concept of a lattice [7],[9]. Gradualness can express a similarity between propositions, levels of uncertainty or preferential degrees.

Relation between the probability theory and fuzzy logic

Zadeh said, in the seminal paper [7], that the concept of fuzzy set is nonstatistical in nature:

" In fact, the notion of a fuzzy set is completely nonstatistical in nature. " [7, p. 340.]

In probability theory and in the theory of fuzzy sets two types of uncertainty are considered: randomness, in the first, and fuzziness – a generalized concept of a classical set, in the latter. Probability deals with mutually disjoint facts, states, or situations. Fuzziness deals with membership degrees to various sets, or phenomena, which are not mutually exclusive. Membership functions may be subjective, i.e. not unique for all observers. Subjectivity, present in a description of membership functions, comes from individual differences in description of abstract concepts, and has a little with randomness. Hence, the subjectivity and nonrandomness of fuzzy sets is the main difference between the study of fuzzy sets and probability theory, which deals with objective treatment of random phenomena.

A probability is defined in a probability space (S, $\mathcal{P}$, P), where $\mathcal{P}$ is a σ-field in S, and P is a probability measure, which maps $\mathcal{P}$ to [0,1]

$$P: \mathcal{P} \to [0,1] \tag{24}$$

A membership function $\mu_A(s)$, $s \in S$, where A is a fuzzy set in S, is a measure

$$\mu_A: S \to [0,1] \tag{25}$$

A probability is defined on the power set $\mathcal{P}(S)$ of the set S. A membership function $\mu_A(s)$ is defined on the basic set, the universe of discourse S. If S is a finite set, for P additivity (10) holds

$$\sum_{s \in S} P(\{s\}) = 1 \tag{26}$$

while μ_A is nonadditive measure:

$$\sum_{s \in S} \mu_A(s) \neq 1 \tag{27}$$

Probability theory is based on probabilistic logic, and the theory of fuzzy sets on fuzzy logic. Probability theory deals with a probabilistic measure, while the theory of fuzzy sets deals with a fuzzy measure [10].

Example 3. A difference between a probability and a membership degree can be illustrated by the next example [11], [10], [1]: Assume S is the set of all liquids. Let A be a fuzzy subset of S, specified by the statement "liquids suitable for drinking ". Pure water has the membership degree to the set A equal to 1, wine has the membership degree to set A equal to 0.6, juice 0.8, brandy 0.2, and hydrochloric acid 0. Suppose a thirsty traveler is in a desert and finds two bottles with liquids. The first bottle is marked by the membership degree to the set of liquids suitable for drinking $\mu_A = 0.92$, and the second by the probability that the liquid is suitable for drinking $P = 0.92$. Which bottle the traveler will choose to drink from? The membership degree $\mu_A = 0.92$ means that in the bottle is something between mineral water and juice, what is quite acceptable. If the choice is the second bottle, the traveler can get pure water, but also poison. The first bottle is an obvious choice.

In the foregoing example, in the terminology of the theory of fuzzy sets, the set *S*, the set of all liquids, is the universe of discourse. The fuzzy subset *A* in the set *S*, is specified by the linguistic statement ″liquids suitable for drinking″. Elements in the set *S* belong to a fuzzy set *A* with some membership degrees, with their values in the domain [0,1].

Thus, probability theory and fuzzy logic model two different phenomena of uncertainty, and they are complementary rather than competitive [14]. Assignment of probabilities to fuzzy events can be considered – probability theory of fuzzy events. Zadeh in [14], points to the fact that uncertainty phenomena exist, where both randomness and fuzziness are present. Namely, the fact that the theory of probability is based on bivalent logic has the consequence that probability theory is lacking in capability to operate with perception-based information, whose validity is a matter of degree. Such information (for example, "The road is wide") has the form of propositions drawn from a natural language, namely beyond the reach of existing predicate-logic-based techniques of meaning representation. Bivalent logic is the logic of measurement, while fuzzy logic is the logic of perception. To enable probability theory to deal also with problems described using perception-based information, Zadeh proposes [14] to restructure probability theory by replacing bivalent logic on which it is based with fuzzy logic. The result of restructuring probability theory is a more general and more complex probability theory, which may be called perception-based probability theory. In standard probability theory only likelihood is a matter of degree, a degree from a unit interval between 0 and 1. In perception-based probability theory, everything – and especially truth and possibility – is, or is allowed to be, a matter of degree. Complexity is the price of constructing a probability theory which has a close rapport with pervasive imprecision, uncertainty and ill-definedness of the real world.

Possibility theory

Possibility theory [12] gives the way of uncertainty modeling when available information is not precise and certain, but is described using fuzzy sets. Possibility theory is connected with the theory of fuzzy sets in such a way that the concept of a possibility distribution is defined as a fuzzy restriction which acts as an elastic constraint on the values that may be assigned to a variable [12]. Similarly to a probability (1 - 4), a possibility can also be considered from the measure theory aspect [27], [6].

Let $A \subset S$, where *A* is a crisp (traditional) set. Let *V* be a variable defined on *S*. To say that *V* takes its value in *A* indicates that any element in *A* could possibly be a value of *V*, and that any element not in *A* cannot be a value of *V*. The statement ″the variable *V* takes its value in *A*″ can be viewed as inducing a possibility distribution π over the set *S*, associating with each element *s* in *S*, the possibility that *s* is a value of *V*

$$\Pi(V=s)=\pi(s)=\begin{cases}1 & if \quad s\in A,\\ 0 & if \quad s\notin A.\end{cases}$$

Next, assume *A* is a fuzzy subset in a universe of discourse $S = \{s\}$, where *s* is a generic element in the set *S*. A fuzzy set *A* is specified by its membership function μ_A, and then a statement of the form " *V* is A", where *V* is a variable taking its values in *S*, induces a possibility distribution Π that makes the possibility that *V* takes value *s* to $\mu_A(s)$. This value, $\mu_A(s)$, represents a compatibility measure of the value *s* with *A*. In that way *V* becomes a fuzzy variable associated with a possibility distribution Π, in much the same manner as a random variable is associated with a probability distribution. A fuzzy set *A* induces a possibility distribution that is equal to a membership function μ_A of that fuzzy set *A*, on the values of *V*

$$\Pi(V=s)=\pi(s)=\mu_A(s)$$

A thesis advanced in [12] is that the imprecision that is intrinsic in natural languages is, in the main, possibilistic rather than probabilistic in nature. Thus, by employing the concept of a possibility distribution, a proposition *p*, in a natural language may be translated into a procedure, which computes the possibility distribution of a set of attributes, implied by *p*. In such a way, in [12] an improvement has been made with respect to the results of Wiener and Shannon: they considered that information was intrinsically statistical in nature. The improvement is related to the fact that when the main concern is with the meaning of information, the proper framework for information analysis is possibilistic rather than probabilistic.

As a probability (1-4), a possibility can also be considered from the point of view of the measure theory (1), (2), (28), (29), [6]:

Definition 5. Let $\mathcal{F}$ be equal to the power set of the set S: $\mathcal{F} = \mathcal{P}(S)$. *A possibility on S* is a belief measure $\Pi: \mathcal{F} \to$ $\to [0,1]$ such that for any system $\{X_\alpha \mid \alpha \in I\}$ of the subsets of *S* (*I* is any non-empty index set) it holds

$$\Pi\left(\bigcup_{\alpha\in I} X_\alpha\right)=\sup_{\alpha\in I}\ \Pi(X_\alpha) \tag{28}$$

A possibility $\Pi(A)$, $A \subset S$, $s \in S$, indicates the possibility that the value of the variable *V*, and that value is *s*, lies in the subset *A*.

A necessity is a belief measure $\mathcal{N}: \mathcal{F} \to [0,1]$ such that for any $\{X_\alpha \mid \alpha \in I\}$ as above

$$\mathcal{N}\left(\bigcap_{\alpha\in I} X_\alpha\right)=\inf_{\alpha\in I}\ \mathcal{N}(X_\alpha) \tag{29}$$

The properties of possibility are given by the following statements [6]:

1. If Π is a possibility on *S* then the function $\mathcal{N}$, defined as $\mathcal{N}(X) = 1 - \Pi(S-X)$, is a necessity measure on *S*. Conversely, each necessity $\mathcal{N}$ defines a possibility $\Pi(X) = = 1 - \mathcal{N}(S-X)$.
2. Each possibility is uniquely determined by its values on singletons: if $\pi: S \to [0,1]$ is a normalized fuzzy subset of *S* (i.e. $\pi(s_i) = \mu(\{s_i\})$ and $\sup_{s\in S} \pi(s) = 1$), then

$$\Pi(X) = \sup_{s\in X} \pi(s) \tag{30}$$

is a possibility. (And, of course, if Π is a possibility and $\pi(s) = \Pi(\{s\})$, then π is a normalized fuzzy subset of *S* determining Π by (30).)

From the foregoing exposition it is obvious that the possibility theory has been derived from the theory of fuzzy sets. A possibility function is connected with the fuzzy set which represents the linguistic variable. For example, if a variable *V* corresponds to the value of a number, and it is known that the value of a number, i.e. the value of a variable V is "*approximately* 1", then the fuzzy set "*approximately* 1" induces the possibility distribution in such a way that the membership degree of an element becomes the possibility of that element.

A fuzzy set and a possibility distribution have a common mathematical expression. But, there is a difference between them: a fuzzy set A can be viewed as a fuzzy value that is assigned to a variable. Viewed as a possibility restriction, A is the fuzzy set of nonfuzzy values that can be possibly assigned to a variable V.

Possibility and probability

Concepts of a probability distribution and a possibility distribution are different [12]. Possibility measures are max-decomposable for the disjunction of events (28), [12], [10], while the probability measures are additive (for mutually exclusive events). Necessity measures are min-decomposable for the conjunction (29). Possibility (respectively, necessity) measures are not compositional for conjunction (respectively, disjunction), eg.

$$\mathcal{N}(A \cup B) \geq \max(\mathcal{N}(A), \mathcal{N}(B)) \tag{31}$$

(One may be (somewhat) certain of A or B without being certain of A or being certain of B, at all.)

In possibility theory, the assessment of uncertainty of A requires two numbers, namely $\Pi(A)$ and $\mathcal{N}(A) = 1 - \Pi(\overline{A})$ which are only weakly related, while the probability of A completely determines the probability of the complementary event $\overline{A}$. Also, possibility (and necessity) measures only require an ordered scale for grading uncertainty, since the possibility theory only uses max, min, and an order-reversing operations. This is in agreement with a rather qualitative view of uncertainty.

Applications of fuzzy sets

The ability of fuzzy sets and possibility theory to model gradual features or elastic constraints the satisfaction of which is a matter of degree, as well as information pervaded with imprecision and uncertainty, makes fuzzy sets useful in a great variety of applications. The most popular area of applications is fuzzy control. The fuzzy set methodology can be used in the area of information systems, especially in information retrieval and database management. If the fuzzy methodology is used, then these systems may allow for the presence of imprecise, uncertain, or vague information in the data base. Problems of fulfillment of elastic constraints and optimization make a class of applications (eg. [30]). Besides that, fuzzy logic, in the context of artificial intelligence, is applied in the following domains:

- knowledge representation;
- approximate reasoning;
- numerical function coding and approximating, in data mining [29];
- pattern recognition and classification;
- multivalued logics;
- processing and analysis of digitalized pictures represented by levels of gray.

Dempster-Shafer theory of evidence

The most important objects in the Dempster-Shafer theory of evidence are Shafer's belief functions. They are used to model and quantify subjective credibility induced in an observer by evidence. Pieces of evidence consist of internal (objective) evidence (facts) and of external evidence (testimonies). Using Dempster's work on lower and upper probabilities [15], Shafer gave [16] a reinterpretation of that work. In that reinterpretation, Dempster's "lower probabilities" are identified as cognitive probabilities or degrees of credibility, the rule of combination of such degrees of credibility is taken as fundamental, and the idea that degrees of credibility arise as lower bounds over classes of Bayesian probabilities is abandoned.

The originality and the power of Shafer's model is that it does not evoke the principle of insufficient reason or an argument of symmetry. In the situations which would give the contradiction, Shafer's model leaves the total mass of belief allocated to the proposition, without splitting it between the components of the proposition [19].

A fuzzy measure μ with respect to a variable V (a measure defined by expressions (1) and (2)), provides a description of a knowledge about the variable. When using a belief measure, although there exists some uncertainty with respect to the actual value of the variable, it is assumed that there exists no uncertainty with respect to the knowledge of the description of the uncertainty. But, the situation is possible in which there exists only partial information with the respect to the underlying belief measure, for example that $\mu(A) \in [a,b]$.

The Dempster-Shafer theory of evidence is an uncertainty modeling tool used as an alternative to the probability theory [6], [20]. The main application is in cases when there is no enough information that should be sufficient for using more precise apparatus of probability theory. The main idea is still probabilistic. Specifically, assume S is the basic set of elementary propositions, and $\mathcal{P}$ the power set of S. A function $g(.)$ is introduced on $\mathcal{P}$ which fulfills the axioms (1) and (2) (the point in $g(.)$ corresponds to the variable of the observed universe of discourse). If p and q denote propositions, the elements from the set $\mathcal{P}$, $\mathcal{P}$ can be treated as Boole's algebra of propositions. The set notation is freely combined with the proposition notation, and then expressions (1) and (2) are

$$g(\varnothing) = 0 \tag{32}$$

$$g(S) = 1 \tag{33}$$

$$\text{if } p \subset q \ (p \Rightarrow q = 1) \text{ then } g(p) \leq g(q); \quad \text{(monotonicity)} \tag{34}$$

where $\varnothing$ is an empty set, S is the whole (basic) set), and "$p \Rightarrow q = 1$" (tautology), means: "it is true that p implies q in any interpretation".

Two standard functions that satisfy those axioms are Shafer's functions of *credibility* and *plausibility*. Previously, the function m, called a *basic probability assignment,* is introduced [6].

Definition 6. Let V be a variable that takes values in the set S. Let $\mathcal{P}$ be the power set of S. *The basic probability assignment* is a function m: $\mathcal{P} \rightarrow [0,1]$, which fulfills the following conditions:

$$m(\varnothing) = 0 \tag{35}$$

$$\sum_{p \in \mathrm{P}} m(p) = 1 \tag{36}$$

The basic probability assignment m defines a probability distribution on the power set $\mathcal{P}$.

If there is a block of insufficiently certain information, it is possible to allocate probabilities to propositions from $\mathcal{P}$, which are not mutually disjoint (in the basic set S elements are always mutually disjoint). The propositions p such that $m(p) > 0$ are called *focal propositions* (because information is based, i.e. ″focused″on them). Certain information is represented by a unique focal proposition p_0 , $m(p_0) = 1$. The value of $m(p)$ is the probability that the evidence is precisely and entirely described by the proposition p. Every basic probability $m(p)$, $p \in \mathcal{P}$, supports also every proposition q that is implied by the proposition p.

Definition 7. *Credibility of a proposition q, $Cr(q)$, is a* sum of all basic probability assignments of propositions p which imply the proposition q (i.e. of all more specific propositions p which are contained in the proposition q, $p \subset q$)

$$\forall q \in \mathcal{P},\ Cr(q) = \sum_{p \Rightarrow q} m(p) \tag{37}$$

The credibility function $Cr(.)$, Cr:$\mathcal{P} \to [0,1]$ has a complementary function, the plausibility function $Pl(.)$: Pl: $\mathcal{P} \to [0,1]$

$$\forall p \in \mathcal{P},\ Pl(p) = 1 - Cr(\neg p) \tag{38}$$

The plausibility of a proposition q is a sum of amounts of belief that are allocated to the proposition p and that are not in contradiction with the proposition q, i.e. which do not imply $\neg q$.

Definition 8. *The plausibility function $Pl(.)$* is a sum of all basic probability assignments of the propositions p for which is fulfilled $\neg(p \Rightarrow \neg q)$, i.e. $p \cap q \neq \emptyset$

$$\forall q \in \mathcal{P},\ Pl(q) = \sum_{\neg(p \Rightarrow \neg q)} m(p) \tag{39}$$

that is, the degree of plausibility of some event q, $Pl(q)$, is the sum of the basic probability assignments of the events p for which is fulfilled $p \cap q \neq \emptyset$.

Hence, it holds

$$\forall p \in \mathcal{P},\ Cr(p) + Cr(\neg p) \leq 1 \tag{40}$$

$$\forall p \in \mathcal{P},\ Pl(p) + Pl(\neg p) \geq 1 \tag{41}$$

$$\forall p \in \mathcal{P},\ Cr(p) \leq Pl(p) \tag{42}$$

Both functions are, by definition, belief functions, they fulfill expressions (32-34). Intuitively, credibility $Cr(p)$ can be interpreted as a degree of minimum or necessary probabilistic support, $P(p)$, to a proposition p, defined on the base of the current, incomplete, knowledge. Plausibility $Pl(p)$ can be interpreted as a degree of maximum or potential support, $P(p)$, to a proposition p.

Two special cases of credibility and plausibility functions are especially interesting, and are specified by the structure of focal elements. These cases are:

1. If every focal proposition p is such that p does not imply q, then p and q are mutually disjoint, $\forall q,\ p \in \mathcal{P}$. In that case the expression for credibility and plausibility are in fact the expressions dealing with probability.
2. If the set of focal propositions can be ordered, then the function Pl is a possibility measure and conversely. In this case Cr is also a necessity measure.

Different pieces of evidence are combined by the application of Dempster's rule of combination on the basic probability assignments. The rule is an associative and commutative binary operation defined on the set S of all basic probability assignments. For two basic probability assignments, m_1 and m_2 , the combined basic probability assignment m_{12} is given as

$$m_{12}(C) = \sum_{A \wedge B = C} (m_1(A) * m_2(B))/(1-d) \tag{43}$$

where

$$d = \sum_{A \wedge B = \emptyset} (m_1(A) * m_2(B)) \tag{44}$$

and A, B, $C \in \mathcal{P}$. The quantity d is the measure of disagreement of two sources, and is used as a renormalization factor.

Since the subset of possibilistic basic probability assignments is not closed with respect to Dempster's rule, the possibilistic rule of combination has been introduced. For this rule the mentioned subset is closed

$$\pi_{12}(s) = \min(\pi_1(s), \pi_2(s)) / (1\text{-}d) \tag{45}$$

where

$$d = \sup_{s \in S} \min(\pi_1(s), \pi_2(s)) \tag{46}$$

For the possibilistic rule of combination, idempotency holds, and that is not case for Dempster's rule or for the multiplication of probabilities. In the case of the possibilistic rule, information sources can be dependent.

Dempster-Shafer's belief functions provide a model for quantifying a degree of belief that the proposition is true. The model is promising for the development of expert systems that need to handle uncertainty. Computer implementations have shown that Dempster-Shafer's approach to uncertainty modeling demands a lot of input data, so computer resources demands are relatively significant [22]. (This approach to uncertainty modeling is used, for example, in the diagnostic expert system GERTS [22]).

Roughness

The theory of rough sets was originated by Zdzislaw Pawlak in the early 1980s as a result of a long-term program of fundamental research on logical properties of information systems [23]. The theory is concerned with the classificatory analysis of imprecise, uncertain or incomplete information or knowledge, expressed in terms of data acquired from experience. The primary notions of the theory of rough sets are [24]: the approximation space and lower and upper approximations of a set. *The approximation space* is a classification of the domain of interest into disjoint categories. The classification formally represents the knowledge about the domain, i.e. the knowledge is understood as an ability to characterize all classes of the classification, for example, in terms of features of objects belonging to the domain. Objects belonging to the same category are not distinguishable, which means that their membership status with respect to an arbitrary subset of the domain may not always be clearly definable. This fact leads to the definition of a set in terms of lower and upper approximations. *The lower approximation* is a description of the domain objects which are known with certainty to belong to the subset of interest. *The upper approximation* is a description of the objects which possibly belong to the subset. *The negative region* is made of those domain objects which are known with certainty not to belong to the subset of interest. Any subset defined through its lower and upper approximation is called a *rough set*.

The previous exposition can be formalized: assume that a domain, a set S of objects, a set of object attributes, AT, a set of values, VAL, and a function $f: S \times \mathrm{AT} \to \mathrm{VAL}$, (that can be given by a table) are given. In that way each object from S is described by the values of its attributes. Then, an equivalence relation (called also *an indiscernibility relation*) $R(\mathrm{A})$ is defined, where $\mathrm{A} \subset \mathrm{AT}$.

Definition 9. If, for any two elements $s_i, s_j \in S$

$$s_i \, R(\mathrm{A}) \, s_j \Leftrightarrow f(s_i, a) = f(s_j, a), \ \forall a \in \mathrm{A}, \ i \neq j \tag{47}$$

then s_i and s_j are *indiscernible elements* with respect to the attributes in A.

The set S is partitioned, using the relation $R(\mathrm{A})$, into equivalence classes. The pair (S, R) forms an *approximation space* with which the arbitrary subsets of S referred to as *concepts* are approximated. *A concept* is a subset of the set S of all objects with the same value of the function f. For a given concept $\Sigma \subset S$, that subset can be approximated by unions of various equivalence classes: by the lower approximation of the set Σ, i.e. by the union of all those equivalence classes in which the set S is partitioned using the relation $R(\mathrm{A})$, and that are the subsets of Σ, or by the upper approximation of the set Σ, i.e. by the union of all those equivalence classes produced by the relation $R(\mathrm{A})$, the intersection of which with the set Σ is not the empty set.

Definition 10. *The lower approximation* of the set Σ, $lower(\Sigma)$, (also known as *the positive region*, $pos(\Sigma)$) is defined by the following expression

$$lower(\Sigma) = pos(\Sigma) = \{s \in S \mid [s]_{\mathrm{R}} \subset \Sigma\} = \bigcup_{[s]_R \subset \Sigma} s \tag{48}$$

where $[s]_R$ denotes the equivalence class of the relation R that contains the elements from the set S with the generic description s.

Definition 11. *The upper approximation* of the set Σ, $upper(\Sigma)$, is defined by the following expression

$$upper(\Sigma) = \{s \in S \mid [s]_{\mathrm{R}} \cap \Sigma \neq 0\} = \bigcup_{[s]_R \cap \Sigma \neq \varnothing} s \tag{49}$$

where $[s]_R$ is as in Defin

A number of practical application of this approach have been reported in the literature [26], in areas as medicine, drug research, process control, and others. One of the advantages of the theory of rough sets is that programs implementing its methods may easily run on parallel computers. An approach that uses the rough set theory is implemented in the empirical learning system called LERS (*Learning from Examples based on Rough Sets*). The LERS has been used for the development of rule-based expert systems applied on the space station *Freedom* [26].

Applicability survey of the considered uncertainty models

On the basis of the previous exposition of different types of uncertainty, in the following table, Table 2, suggestions are given on applicability of the considered uncertainty models:

Table 2. Applicability of uncertainty models

	ES	L	DM	DR	Cl	FA	PP
P	+	+	+	+	+		
CF	+						
F	+	+	+	+	+	+	+
Po	+	+					
DS	+	+					
R	+	+	+	+	+	+	+

The symbols in Table 2 are: P-probability, CF- certainty factors, F-fuzziness, Po – possibility, DS- Dempster-Shafer, R-roughness, ES-expert systems, L-learning, DM-data mining, DR- data reduction, Cl-classification, FA- function approximation, PP-picture processing.

The starting point for a decision about the uncertainty modeling formalism in an application problem is the set of available data and the kind of uncertainty seen in those data: if the source of uncertainty is a linguistical uncertainty, then the applied formalism should be the one of fuzziness, which possesses the possibility of processing that kind of uncertainty. In the expert system application context, the quality of a reasoning model inherent to the uncertainty should be considered as well. Practical issues of knowledge elicitation and representation from given data using the uncertainty model, are also important, as well as the applicability of the uncertainty paradigm in the decision-making process.

Conclusion

In the paper it is pointed to three different types of uncertainty. The relationship between those uncertainty types has been shown. The short survey of uncertainty modeling formalisms applied in expert systems is given. The definitions and the survey of the basic features of the basic uncertainty types are given: for randomness (probability, certainty factors, Dempster-Shafer's theory), for fuzziness (fuzzy set theory, possibility theory), and for roughness (the theory of rough sets). Some others modeling formalisms, such as nonmonotonic logics and default reasoning, and qualitative versions of probability, such as the Spohn calculus and kappa-calculus, have not been considered here due to controversies about these approaches. The central point in the exposition is the notion of the belief measure, and that notion can be used for unified uncertainty modeling in cases of randomness and fuzziness. Roughness is defined on the basis, not of the function, but of characteristic sets: the classification region, the lower and the upper approximations. The difference between probability and fu-ziess, expressed also in a formal mathematical manner, is shown. The discussed formalisms of uncertainty modeling are not concurrent, but complementary, they describe different types of uncertainty.

The conditions have been defined for the application of the discussed approaches to uncertainty modeling, which makes easier the application of that modeling in expert and other artificial intelligence systems. Probability theory can be viewed as a generalization of classical propositional logic that is useful when the truth of particular proposition is uncertain. The main form of probabilistic inference is to use Bayes's theorem to go from a prior probability on a proposition to a posterior probability conditioned on a new evidence. A prior probability distribution must be given over the propositions of interest. The principle of insufficient reasoning should be used to assign these initial (prior) probabilities. Probability is mathematically well-founded, precise, but requests a large data sample for correct results.

Fuzzy sets and possibility theory are useful in applications characterized by existence of gradual properties of objects, soft constraints the satisfaction of which is a matter of degree, as well as information pervaded with imprecision and uncertainty. Fuzzy approach is used when there is a need for models capable of handling the kind of knowledge that humans manipulate with (i.e. ill-defined classes, classes with imprecisely located boundaries, classes with gradual membership and non-probabilistic uncertainty, vague predicates, imprecise or uncertain information, expert rules pervaded with vagueness or exceptions). Also, the fuzzy approach is used when a tool is needed for representing and reasoning with the available information in a manner similar to the way humans express knowledge and summarize data.

When there is no any preliminary or additional information about data, such as probability distribution in statistics, basic probability assignments in the Dempster-Shafer theory, or grade of membership or the value of possibility in the fuzzy set theory, the rough set theory may be used. It uses only information given by the operationalized data, and does not rely on other model assumptions. Imprecision is expressed by quantitative concepts (approximations). The rough set theory requests that granularity of the domain of interest can be expressed by partitions and their associated relations on the set of objects.

Practical reasoning requires some schemes for representing uncertainty. Many real-world applications of probabilistic methods exist, and fuzzy logic and the rough sets theory have, notable success as well. Each method has its merits, and may be suitable for practical applications.

Reports in references and some practical experience [30] show that implementations of uncertainty models using certainty factors and fuzziness are efficient, with linear evaluation time. Probabilistic reasoning is a nondeterministic polynomial, i.e. potentially inapplicable for massive data. The computer implementations of Dempster-Shafer's model are more complex than those of a simple probabilistic model. Dempster-Shafer's model requests massive input data, demands for computer resources are considerable.

Abilities to process uncertain information and to reason on the basis of insufficient knowledge are determining features of intelligent behavior in an *uncertain*, i.e. complex and dynamical, environment. The uncertainty modeling is the important research field in the domain of expert systems.

In the area of uncertainty modeling there are many research problems, which remain to be solved, either in the formal-mathematical part of the area, or in its aspects related to applications. Many theoretical problems still await proper clarification especially in the areas fuzzy set theory and the theory of rough sets.

The possibilities of further research work exist, either in connection with some type of uncertainty, or in connection with those uncertainty phenomena in which more types of uncertainty are present. This research work could encompass, for example, the restructuring of probability theory by replacing bivalent logic on which it is based with fuzzy logic. The result should be a more general and more complex probability theory, perception-based probability theory, capable to deal with problems that in their description contain information based on perception. Besides that, this research work should encompass open problems from the theory of fuzzy sets, for example the problem of choosing operators in connection with mathematical foundations of fuzzy logic, various problems of fuzzy logic applications in medicine, control, with databases, in search machines on the Internet, and many other problems. Fuzzy controllers are a very successful area of applications of fuzzy logic. Rough controllers, i.e. controllers based on the rough set theory, and even probabilistic controllers, i.e. controllers based on probability theory, seem to be very promising areas of applications.

References

[1] ŠALETIĆ,D.Z., VELAŠEVIĆ,D.M. Uncertainty treating formalisms based on fuzzy sets, in Dr. D. RADOJEVIĆ, ed.: *Proceeding of abstracts of SOFT COMPUTING DAY*, pp. 13, XXV Yugoslav symposium on operations research, SYM-OP-IS '98., Herceg-Novi, 22.9.1998. (in Serbian)

[2] ŠALETIĆ,D.Z. *On Expert Systems and Their Military Applications.* Scientific-technical information 8/86, Military-Technical Institute, Belgrade, Yugoslavia, 1986. (in Serbian)

[3] DURKIN,J. *Expert Systems.* Macmillan Publishing Company, New York, 1994.

[4] ŠALETIĆ,D.Z., VELAŠEVIĆ.D.M. *Formalizmi modeliranja neizvesnosti u ekspertskim sistemima u neizvesnim okruženjima.* Zbornik radova VIII simpozijuma o računarskim naukama i informacionim tehnologijama YiNFO 2002, CDROM, Kopaonik, 11-15.03.2002., pp.136.pdf, 6 p.

[5] ELKAN,C. The Paradoxical Success of Fuzzy Logic. *IEEE Expert* – A Fuzzy Logic Symposium, 1994, vol.9, no.4, pp.3-8.

[6] HÁJEK,P. *Metamathematics of Fuzzy Logic.* Kluwer Academic Publishers, Dordrecht, 1998.

[7] ZADEH,L.A. Fuzzy Sets. *Information and Control*, 1965, vol.8, no.33, pp.338-353.

[8] Fuzzy Mail Digest No. 481, http://www.dbai.tuwien.ac.at/marchives/fuzzy-mail/index.html, [ZADEH,L.A.: Berkley Initiative in Soft Computing (BISC) Seminar: Fuzzy Sets and Fuzzy Logic: Discovery and Controversy, Abstract (bisc-group@berkley.edu)]

[9] KAUFMANN,A. *Introduction to the Theory of Fuzzy Subsets*, vol.1, Academic Press, New York, 1975.

[10] PAP,E. *Fuzzy Measures and Their Applications*, University in Novi Sad, Mathematical faculty, Novi Sad, F.R. Yugoslavia, 1999. (in Serbian)

[11] BEZDEK,J. Fuzzy Models – What Are They, and Why?, *IEEE Transactions on Fuzzy Systems*, February 1993, vol.1, no.1, pp.1-6.

[12] ZADEH,L.A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1978, no.1, pp.3-28.

[13] SUGENO,M. Fuzzy Measures and Fuzzy Integrals – A Survey, in Gupta,M.M., Saridis,G.N., Gaines,B.R. *Fuzzy Automata and Decision Processes,* North Holland, New York, 1977, pp.89-102.

[14] ZADEH,L.A. Probability Theory and Fuzzy Logic, http://ins.cnu.nctu.edu.tw/JAN2003/Probability Theory and Fuzzy Logic.pdf, Decembar 2002.

[15] DEMPSTER,A.P. Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Statist.,* 1967, no.38, pp.325-339.

[16] SHAFER,G. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[17] DUBOIS,D., PRADE,H. A Set-Theoretic View of Belief Functions, *Int. J. General Systems,* 1986., vol.12, pp.193-226.

[18] DUBOIS,D., PRADE,H. Treating of Uncertainty in Knowledge-Based Systems Using Fuzzy Sets and Possibilistic Theory. *Int. J. of Intelligent Systems,* 1988., vol.3, pp.141-165.

[19] SMETS,P. Belief Functions, in SMETS,P., MAMDANI,E., DUBOIS,D., PRADE,H. (ed.) *Non-Standard Logics and Automated Reasoning,* Academic Press, London, 1988, pp.253-286.

[20] BOŽINOVIĆ,R., KRTOLICA,R., OBRADOVIĆ,I. Interpretation of Results of Statistical Tests Using the Dempster-Shafer theory, *Proc. XV Yugoslav Symposium on Operations Research*, SYM-OP-IS '88, Brioni, S.F.R.Yugoslavia, 1988., pp.3-6. (in Serbian).

[21] YAGER,R.R., KACPRZYK,J., FEDRIZZI,M. *Advances in the Dempster-Shafer Theory of Evidence.* John Wiley, New York, 1993.

[22] YEN,J. A Dempster-Shafer Approach to Diagnosis Hierarchical Hypothesis. *Comm. of the ACM*, May 1989, vol.32.,no.5, pp.573-585.

[23] PAWLAK,Z. Rough Sets. *Intern. J. Inform. Comput. Sci*., 1982, vol.11, no.5, pp.341-356.

[24] A Brief Introduction to Rough Sets. *Electronic Bulletin of Rough Set Community*, EBRSC, 1993.

[25] PAWLAK,Z. Rough sets and fuzzy sets. *Fuzzy Sets and Systems*, 1985, no.17, pp.99-102.

[26] PAWLAK,Z., GRZYMALA-BUSSE,J., SLOWINSKI,R., ZIARKO,W. Rough sets. *Comm. of the ACM*, 1995, vol.38, no.11, pp.89-95.

[27] YAGER,R.R. Uncertainty Representation Using Fuzzy Measures, *IEEE Trans. on Systems, Man and Cybernetics Part B – Cybernetics*, 2002, vol.32, no.1, pp.13-20.

[28] JOHNSON,L., KERAVNAU,E.T. *Expert Systems Technology, A Guide.* Abacus Press, Cambridge Massachusetts, 1985.

[29] ŠALETIĆ,D.Z., VELAŠEVIĆ,D.M., MASTORAKIS,N.E. Comparative Analysis of Some Objective Function Based Fuzzy Clustering Algorithms Applied in Fuzzy Systems Modeling, in MASTORAKIS,N.E., ed.: *Advances in Scientific Computing, Computational Intelligence and Applications*, World Scientific and Engineering Society Press, 2001, pp.333-338, ISBN: 960-8052-36-X

[30] ŠALETIĆ,D.Z., VELAŠEVIĆ,D.M. Missile System Selection Based on the Fuzzy Set Theory. *Yugoslav Journal of Operations Research*, 2000, vol.10, no.1, pp.47-61.

Received: 3.3.2003

Pojavni oblici neizvesnosti i formalizmi njihovog modeliranja u ekspertskim sistemima

U radu je ukazano na različite pojavne oblike neizvesnosti i na odnose između njih. Dat je kratak pregled matematičkih formalizama modeliranja neizvesnosti u ekspertskim sistemima. Pokazana je razlika, iskazana i formalno-matematički, između verovatnoće i rasplinutosti. Definisani su uslovi koji određuju primenu pojedinih razmatranih pristupa modeliranju, čime se olakšava primenljivost modeliranja neizvesnosti u ekspertskim i drugim sistemima veštačke inteligencije. Razmatrane su mogućnosti daljeg istraživačkog rada u tretiranoj oblasti.

Ključne reči: **računarska tehnika, veštačka inteligencija, ekspertski sistemi, neizvesnost, rasplinuta logika, verovatnoća, teorija grubih skupova.**

Phénomènes d'incertitude et les formalismes de leur modélisation chez les systèmes experts

Les phénomènes différentes d'incertitude et leurs relations mutuelles sont traitées. Les formalismes mathématiques de leur modélisation chez les systèmes experts sont brièvement donnés. La différence entre la probabilité et le flou, exprimée en termes formels et mathématiques, est démontrée. Les conditions déterminant l'application de quelques approches à la modélisation sont définies, ce qui facilite l'applicabilité de l'incertitude chez les systèmes experts et d'autres systèmes de l'intelligence artificielle. Les possibilités des recherches plus approfondies dans ce domaine sont considerées.

Mots-clés: **informatique, intelligence artificielle, systèmes experts, incertitude, logique floue, probabilité, théorie des ensembles rudes.**