Control and Cybernetics

vol. **34** (2005) No. 2

On practical problems with the explanation of the difference between possibility and probability

by

Andrzej Piegat

Faculty of Computer Science and Information Systems
Technical University of Szczecin
Żołnierska 49, 71-210 Szczecin, Poland
e-mail: Andrzej.Piegat@wi.ps.pl

Abstract: In his famous paper "Fuzzy Sets as a Basis for a Theory of Possibility" (Zadeh, 1978) Professor Lofti Zadeh introduced the notion of possibility distribution π_X and the concept of possibility measure. He defined in the paper the possibility distribution function to be numerically equal to the membership function ($\pi_X = \mu_F$). In this paper Professor Zadeh draws the special attention of the reader to the fact that: " ... there is a fundamental difference between probability and possibility". To explain this difference he had given a special example illustrating the difference, which then was cited by many authors of books on Fuzzy Set Theory and gained great importance for understanding the notion of possibility. In the paper the author presents his doubts as to this important example, explains why it is incorrect and gives a correct version of the example based on the notion of possibility distribution of Dubois and Prade.

Keywords: fuzzy systems, fuzzy arithmetic, possibility, probability

1. Introduction

The example of Professor Zadeh's, explaining the difference between possibility and probability given in Zadeh (1978) has contents as below.

"To illustrate the difference between probability and possibility by a simple example, consider the statement 'Hans ate X eggs for breakfast', with X taking values in $U = \{1, 2, 3, 4, \ldots\}$. We may associate a possibility distribution with X by interpreting $\pi_X(u)$ as the degree of ease with which Hans can eat u eggs. We may also associate a probability distribution with X by interpreting $p_X(u)$ as the probability of Hans eating u eggs for breakfast. Assuming that we employ

Table 1. The possibility and probability distributions associated with X

u	1	2	3	4	5	6	7	8
$\pi_X(u)$	1	1	1	1	0.8	0.6	0.4	0.2
$p_X(u)$	0.1	0.8	0.1	0	0	0	0	0

some explicit or implicit criterion for assessing the degree with which Hans can eat u eggs for breakfast, the values of $\pi_X(u)$ and $p_X(u)$ might be as shown in Table 1.

We observe that, whereas the possibility that Hans may eat three eggs for breakfast is 1, the probability that he may do so might be quite small, e.g., 0.1. Thus, a high degree of possibility does not imply a high degree of probability, nor does a low degree of probability imply a low degree of possibility. However, if an event is impossible, it is bound to be improbable."

This example of Professor Zadeh, illustrating and explaining the difference between possibility and probability is universally cited by many authors of books and other publications on Fuzzy Set Theory. It is of a very great importance for the way we understand the notion of possibility. Sometimes, in the publications, the role of eggs is played by cups of tea or coffee. The example of Professor Zadeh is quoted by e.g. H.J. Zimmermann in his known and many times republished book "Fuzzy Set Theory And Its Applications" (Zimmermann, 1996), where it is labeled an "impressive example". The example is also cited in the known book of Driankov, Hellendoorn and Reinfrank (1993), and in at least two books of Polish authors. Professor Zadeh's example seems to the present author to be not correct and his doubts will be presented below.

In Zadeh (1978) Professor Zadeh introduced certain notation: "Let X be a variable which takes values in a universe of discourse U, with the generic element of U denoted by u and $X = u$ signifying that X is assigned the value u, $u \in U$."

Fig. 1 shows the probability distribution of the event "Hans **eats** u eggs for breakfast" and the possibility distribution understood as "the degree of ease with which Hans **can eat** u eggs" resulting from Table 1 given by Professor Zadeh.

2. Remarks of the author

The probability distribution $p_X(u)$ in Fig. 1 refers to the probability of the event "Hans eats u eggs for breakfast" and more precisely "Hans eats u eggs for breakfast according to his appetite". In this event Hans does not force him to eat more than he wants to eat. Let the variable determining the egg number, which Hans eats for breakfast be denoted as u_1 and its universe of discourse as U_1. As can be concluded from Fig. 1 Hans does not eat more than three eggs for breakfast and therefore the universe of discourse can be assumed as

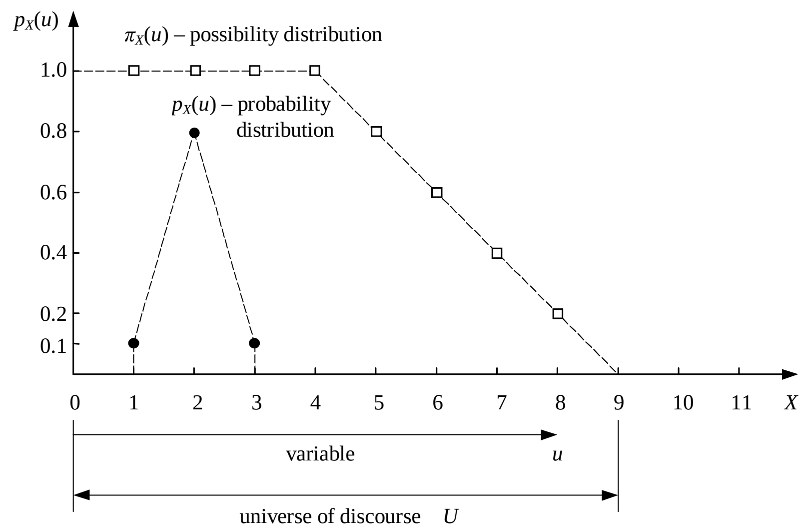

Figure 1. Probability distribution $p_X(u)$ of the event "Hans **eats** u eggs for breakfast" and possibility distribution $\pi_X(u)$ of "the degree of ease with which Hans **can eat** u eggs"

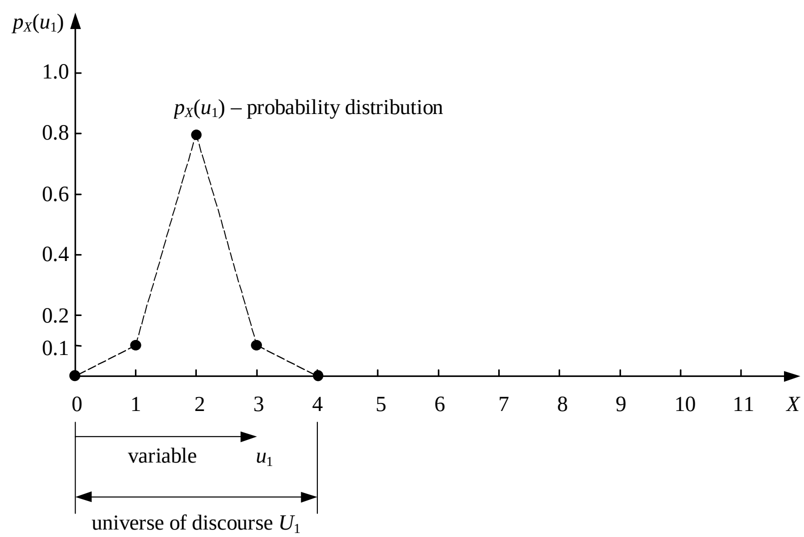

Figure 2. Probability distribution $p_X(u_1)$ of the egg number, which Hans eats for breakfast according to his appetite

$U_1 = \{0, 1, 2, 3, 4\}$, Fig. 2 (the universe was broadened to show elements having probability equal to 0).

The possibility distribution $\pi_X(u)$ in Fig. 1 refers to the variable "number of eggs which Hans **can eat** independently of his appetite". It is another variable than the variable u_1 referring to the probability distribution $p_X(u_1)$ and therefore it will be denoted u_2. If Hans eats with appetite maximally three eggs, then to eat eight eggs (Fig. 1) he eats the successive eggs with smaller and smaller appetite. In particular, to consume the last eggs he must strongly force himself. Let us notice in Fig. 1 that Hans is not able to eat nine eggs.

The variable u_2 (number of eggs which Hans **can** (is able to) **eat**) has a different universe of discourse U_2 than the variable u_1, Fig. 3.

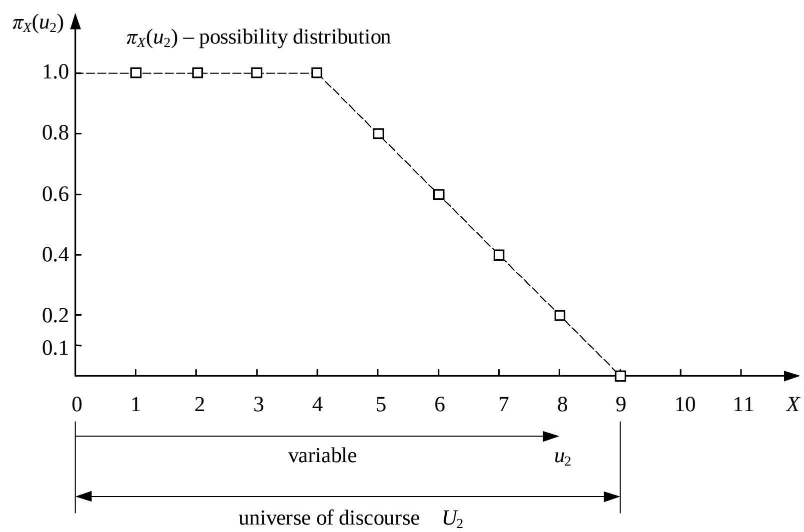

Figure 3. Possibility distribution $\pi_X(u_2)$ of the number of eggs which Hans can (is able to) eat, even when he must force him to eat the eggs

The fact that the both variables u_1 and u_2 are expressed by the "egg number" does not mean that they are identical variables, similarly as the fact that income and costs of a firm are expressed in e.g. Euro does not mean that income and costs is the same.

REMARK 2.1 *The variable u_1 (number of eggs which Hans* **ate** *for breakfast) associated with the probability distribution $p_X(u_1)$ is a different variable from the variable u_2 (number of eggs which Hans* **can** *(is able to)* **eat***) associated with the possibility distribution $\pi_X(u_2)$. Both variables describe different events and have different universes of discourse. Therefore the comparison made by Professor Zadeh seems not correct. A correct comparison of probability and* *possibility should refer to the same and identical variable, e.g. to u_1 (number of eggs, which Hans* **ate** *for breakfast) or to u_2 (number of eggs that Hans* **can** *(is able to)* **eat** *for breakfast). For one and the same variable a probability distribution $p_X(u)$ and a possibility distribution $\pi_X(u)$ should be determined and a comparison of both distributions should be made.*

REMARK 2.2 *The example of Professor Zadeh suggests that possibility distribution is connected with the word "***can***" occurring in description of the event represented by variable X. This leads the reader into error and is not true because possibility distribution can be determined for variables representing any event and not only containing in their description the word* **can** *or a similar expression. It will be shown below that possibility distribution can be determined for the variable "egg number, which Hans ate for breakfast (according to his appetite)" which does not contain the word* **can***.*

REMARK 2.3 *The example of Professor Zadeh does not show that possibility distribution should be used as mathematical description of the event $(X = u)$ only then, when measurements of variable X are inaccurate (uncertain) and when they are of special character – when they are nested measurements (Dubois, 1988; Klir, 1988; Zimmermann, 1996). If measurements of variable X are precise (e.g. egg number eaten by Hans for breakfast can be very easily exactly measured), then possibility distribution must not be used as mathematical description of the event $(X = u)$ because it is a poor information (an example will prove this statement). If measurements of X are precise we may use other, more informatively valuable forms of mathematical description like, e.g., probability distributions.*

REMARK 2.4 *In his example Professor Zadeh did not show that possibility distribution should always be used together with its dual counterpart, with necessity distribution. Only both distributions constitute the complete mathematical description of the event occurrence. Description with only possibility distribution is incomplete.*

3. Short reminder of the scientific concepts of possibility, necessity and probability

There exist various interpretations of the notion possibility. Examples of the interpretations can be found, e.g., in Borgelt and Kruse (2003), Ruspini (1990), Spohn (1990), Spott (1997). In this paper we will use the interpretation of very well known specialists of fuzzy sets: D. Dubois, H. Prade, G. Klir, which can be found in Dubois and Prade (1988, 1993, 1994, 2002), Dubois et al. (2004), Klir and Folgert (1988). This interpretation seems to be dominating and the most approved one and has been published many times between 1988 and 2004 in scientific journals. It allows for determining a possibility distribution from measurements. This interpretation will be shortly explained below.

Let us denote by U the domain of the variable X, by A an event $A \subseteq U$, by $\pi(u)$ the possibility distribution $\pi(u) = \Pi(\{u\})$, where π is a mapping of U into $[0, 1]$. When the set U is finite, every possibility measure Π can be defined in terms of its values on singletons of U (Dubois and Prade, 1988), formula (1):

$$\forall A, \qquad \Pi(A) = \sup\{\pi(u) \mid u \in A\} \tag{1}$$

where $\Pi(A)$ is a possibility measure of occurrence of the event $A \subseteq U$. The notion of possibility $\Pi(A)$ of the event A occurrence and the connected notion of necessity $N(A)$ of the event occurrence can be used only when focal elements E_i of a body of evidence $(\mathcal{F}, m)$ are nested, which means that the focal elements satisfy the condition $E_1 \subset E_2 \subset \cdots \subset E_p$, (Klir and Folgert,1988). Necessity measure $N(A)$ of an event A as a dual measure in relation to possibility measure can be calculated with formula (2) (Klir and Folgert, 1988):

$$\begin{aligned} N(A) &= 1 - \Pi(\bar{A}), \\ N(A) &= \inf\{1 - \pi(u) \mid u \notin A\}, \end{aligned} \tag{2}$$

where $\bar{A}$ is the complement of A.

$\Pi(A) = 1$ means that A is fully possible. The dual measures of possibility $\Pi(A)$ and necessity $N(A)$ are connected by relations (3), (Dubois and Prade, 1988):

$$\forall A \subseteq U, \qquad \begin{array}{lll} N(A) > 0 & \Rightarrow & \Pi(A) = 1 \\ \Pi(A) < 1 & \Rightarrow & N(A) = 0 \\ \Pi(A) \geq N(A) & & \\ N(A) + N(\bar{A}) \leq 1 & & \\ \Pi(A) + \Pi(\bar{A}) \geq 1. & & \end{array} \tag{3}$$

As it results from formulas (1)-(3) to calculate the possibility measure $\Pi(A)$ or the necessity measure $N(A)$ of the event A occurrence the possibility distribution $\pi(u)$ is necessary. The possibility distribution can be determined with formula (4), (Dubois and Prade, 1988):

$$\forall u, \qquad \pi(u) = \begin{cases} \displaystyle\sum_{j=i}^{p} m(E_j) & \text{if} \quad u \in E_i\,, \\ 0 & \text{if} \quad u \in U - E_p\,, \end{cases} \tag{4}$$

where $m(E_j)$ are probability masses carried by the nested subsets E_1, E_2, …, E_p of the universe U (presumed finite) satisfying conditions (5):

$$\begin{aligned} &\sum_{i=1}^{p} m(E_i) = 1 \\ &\forall i, \quad m(E_i) > 0 \end{aligned} \tag{5}$$

"The probability mass $m(E_i)$ can be interpreted as a global allocation of probability to the whole set of elementary events making up E_i, without specifying how this mass is distributed over elementary events themselves" (Dubois and Prade, 1988). Let us denote the measurement (observation data) by $\{I_k \mid k = 1, \ldots, q\}$. **Each measurement I_k is assigned uniquely to the smallest reference set E_i capable of including it.** The probability masses can be calculated with formula (6):

$$\forall i, \quad m(E_i) = \frac{1}{q} \text{ [number of measurements assigned to } E_i \text{]}. \tag{6}$$

If we have the possibility distribution $\pi(u)$ then we can determine the probability distribution $p(u)$ according to the method described below (after Dubois and Prade, 1988). Given a possibility measure in the form of nested focal elements with probability weightings, we may seek to approximate to it by means of probability measure, by interpreting each focal element E_i as a conditional probability $P(\cdot|E_i)$ uniformly distributed over E_i. The atom of probability associated with the element $u \in U$ (finite) is then:

$$\forall u, \quad p(u) = \sum_{i=1}^{P} (u \mid E_i)\, m(E_i) = \sum_{u \in E_i} \frac{m(E_i)}{|E_i|}\ , \tag{7}$$

where $|E_i|$ is the number of elements in E_i. We have therefore made a choice (which could be considered somewhat arbitrary) of one probability measure in the class of all those that satisfy the inequalities (8):

$$\forall A, \quad N(A) \leq P(A) \leq \Pi(A). \tag{8}$$

The probability atoms $\{p(u_i) \mid i = 1, \ldots, n\}$ can be calculated directly from the possibility distribution $\{\pi(u_i) \mid i = 1, \ldots, n\}$:

$$p(u_i) = \sum_{j=i}^{n} \frac{1}{j} \{\pi(u_j) - \pi(u_{j+1})\}\ , \tag{9}$$

where $\pi(u_1) = 1 \geq \pi(u_2) \geq \cdots \geq \pi(u_{n+1}) = 0$, and u_{n+1} is a dummy element (U has n elements). It can readily be seen that (9) defines a one-to-one correspondence between the distributions p and π. The inverse of this formula is (10):

$$\pi(u_i) = \sum_{j=i}^{n} \min(p(u_i), p(u_j)) \tag{10}$$

4. Example illustrating the difference between possibility and probability of an event

In Professor Zadeh's example the difference between possibility and probability of the event occurrence is explained on example of the event "Hans ate X eggs for breakfast". Probability distribution of this event is given in Table 1 and in Fig. 2. In order to clearly explain the difference between probability and possibility of the event occurrence the way how the probability distribution $p_X(u)$ and the possibility distribution $\pi_X(u)$ is determined should be shown and compared.

4.1. The way to determine the probability distribution $p_X(u)$ of the event "Hans ate X eggs for breakfast"

To determine this distribution the information, e.g. from Hans' mother, about the number of eggs, which Hans ate for breakfast, is necessary. Let us assume we have the information as given below:

a) For 1 of all 10 observed breakfasts Hans ate 1 egg.
b) For 8 of the 10 observed breakfasts Hans ate 2 eggs.
c) For 1 of the 10 observed breakfasts Hans ate 3 eggs.

On the basis of the above information we can calculate probability of particular events, formula (11) and Fig 4.

$$\begin{aligned} p_X(u=1) &= 0.1 \\ p_X(u=2) &= 0.8 \\ p_X(u=3) &= 0.1 \end{aligned} \tag{11}$$

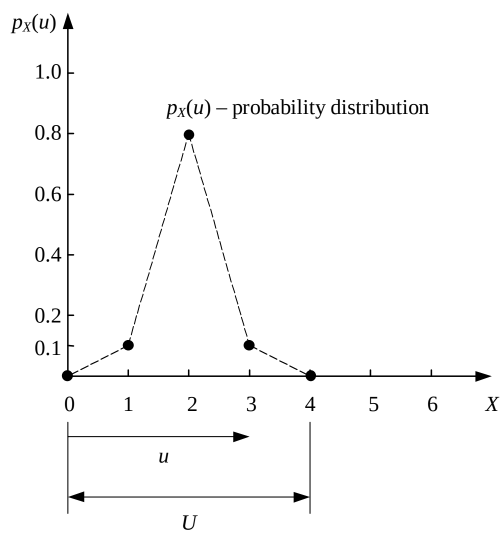

Figure 4. Probability distribution $p_X(u)$ of the event "Hans ate X eggs for breakfast"

Remark 4.1 *In the example a very small number of 10 observations was assumed, which does not satisfy the requirements of the probability theory. However, the author assumed such small number explain easier the problem.*

Remark 4.2 *The evidence information delivered by Hans' mother is precise and refers to each separate number 1, 2, 3 of eggs.*

Remark 4.3 *There exists only one probability distribution $p_X(u)$ resulting from the evidence information a), b), c) delivered by the mother.*

4.2. The way to determine the possibility distribution $\pi_X(u)$ of the event "Hans ate X eggs for breakfast"

Let us assume that we did not obtain from Hans' mother as precise information as a), b), c) from Section 4.1, but a less precise information as follows:

A) It is certain that for 8 of all 10 observed breakfasts Hans ate 2 eggs. (Remark: the above information does not exclude the possibility that Hans ate 2 eggs for 9 or even for all 10 breakfasts.)

B) It is certain that for all 10 observed breakfasts the number of eggs eaten by Hans was in the interval 1 – 3. (Remark: the above information does not necessarily mean that Hans certainly ate for at least 1 breakfast 1 or 3 eggs. It is further possible that Hans could eat e.g. 2 eggs for all 10 observed breakfasts.)

Information "A" defines the focal element $E_1 = \{2\}$. 8 measurements (observations) I_k from the full measurement set $\{I_k \mid k = 1, \ldots, 10\}$ are uniquely assigned to this element. Information "B" defines the focal element $E_2 = \{1, 2, 3\}$ to which are assigned all 10 measurements I_k of the egg number. Now, with formula (6) the probability masses $m(E_i)$ associated with particular focal elements E_i can be calculated:

$$
\begin{aligned}
&m(E_1) = 8/10 = 0.8\\
&m(E_2) = (10 - 8)/10 = 0.2\\
&m(E_1) + m(E_2) = 1.
\end{aligned}
\tag{12}
$$

Probability mass $m(E_{i+1})$ means the probability increase of the event ($X = u$)-occurrence caused by extension of the interval of possible values of the variable u of the focal element E_{i+1} in relation to the focal element E_i nested in it. Fig. 5 shows the focal elements E_i resulting from information "A" and "B" delivered by Hans' mother.

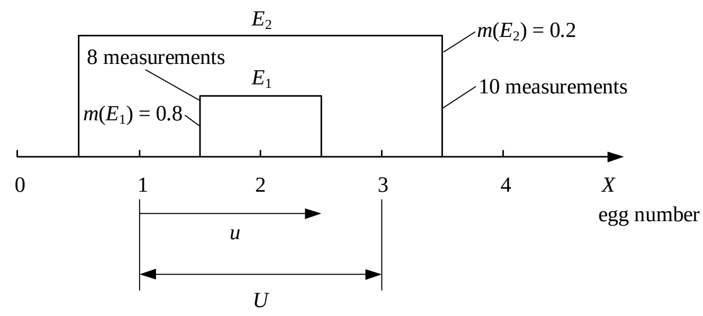

Figure 5. Nested focal elements E_i and probability masses $m(E_i)$ associated with them, which create the body for evidence for the problem of the number for eggs eaten by Hans for breakfasts

With formula (4) the possibilities $\pi_X(u)$ of the event "Hans ate X eggs for breakfast" or $(X = u)$ can now be calculated:

$$\begin{aligned} \pi(u=1) &= \sum_{j=2}^{2} m(E_j) = m(E_2) = 0.2 \\ \pi(u=2) &= \sum_{j=1}^{2} m(E_j) = m(E_1) + m(E_2) = 0.8 + 0.2 = 1 \\ \pi(u=3) &= \sum_{j=2}^{2} m(E_j) = m(E_2) = 0.2\,. \end{aligned} \tag{13}$$

Fig. 6 presents the obtained possibility distribution $\pi_X(u)$.

Let us consider now what is the meaning of the sentence: "Possibility of the event 'Hans ate 1 egg for breakfast' equals 0.2". From information "A" we know that Hans ate at least for 8 from 10 breakfasts two eggs. From information "B" we know that for all 10 breakfasts the egg number did not go outside the interval 1-3. So, we can conclude that number of breakfasts, for which Hans ate only 1 egg could be at most 2, but it could also be 1 or 0. The same refers to 3 eggs. There exist 6 possible distributions of the number of breakfasts for which Hans ate u eggs, Fig. 7.

Because information "A" and "B" delivered by Hans' mother is not precise we do not know which of all possible distributions a-f, Fig. 7, really occurred. Each distribution of the number of breakfasts for which Hans ate $(X = u)$ eggs generates one specific probability distribution $p_{Xi}(u)$, Fig. 8.

Let us notice that the greatest possible probability that Hans ate 1 egg for breakfast results from the distribution $p_{X1}(u)$ in Fig. 8a and is just equal to 0.2. It also means that the **possibility of occurrence of the event** $(u = 1)$ **equals** $\pi_X(1) = 0.2$.

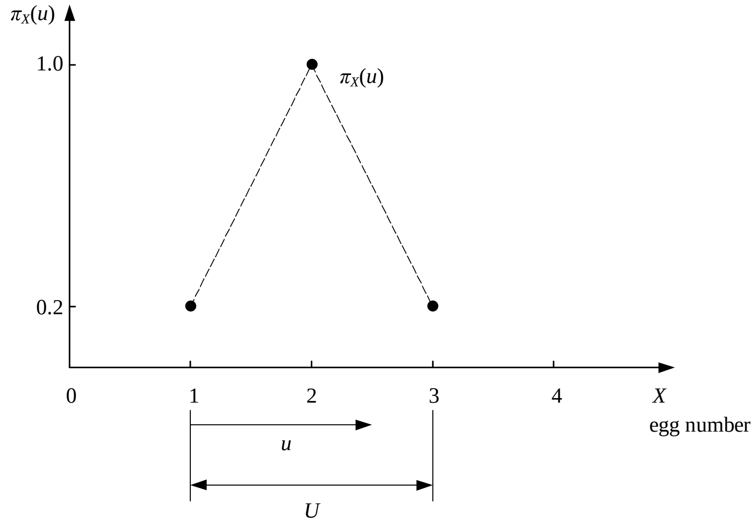

Figure 6. Possibility distribution $\pi_X(u)$ of the occurrence of event $(X = u)$: "Hans ate X eggs for breakfast"

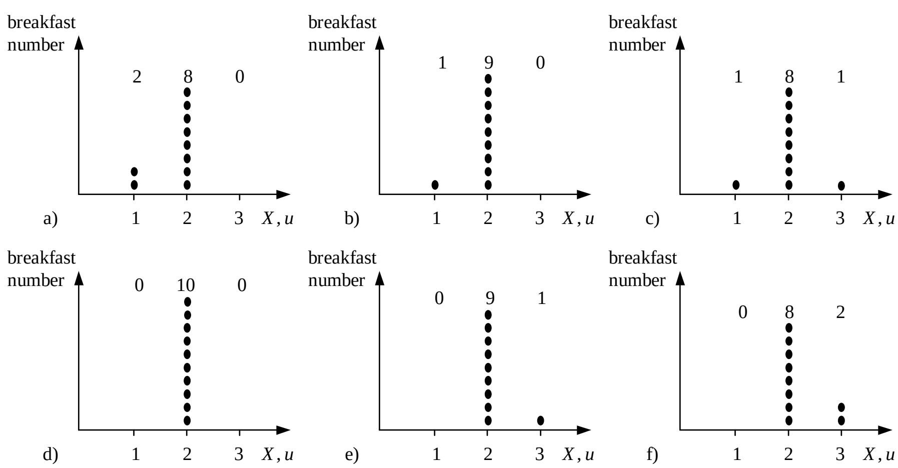

Figure 7. Possible distributions of the number of breakfasts, for which Hans ate $(X = u)$ eggs

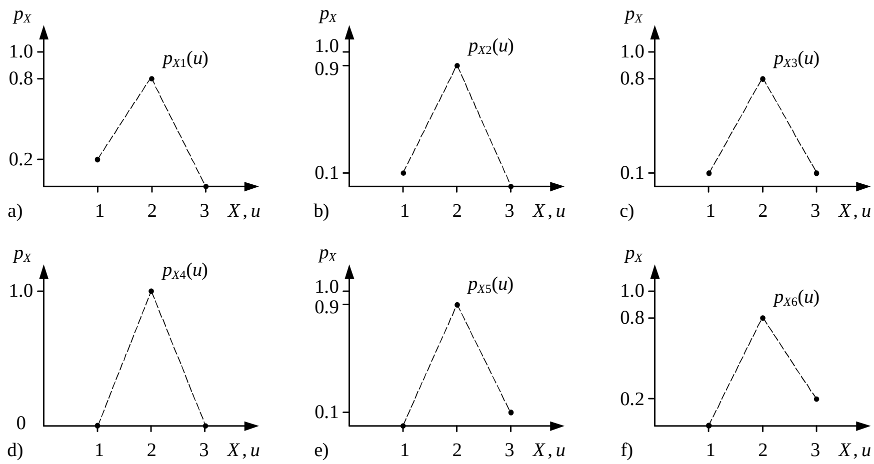

Figure 8. Six feasible probability distributions $p_{Xi}(u)$ of the event $(X = u)$ – "Hans ate X eggs for breakfast" resulting from the inexact information "A" and "B" delivered by Hans' mother

The greatest possible probability that Hans ate two eggs for breakfast results from the probability distribution $p_{X4}(u)$ in Fig. 8d and it equals 1. It means that the possibility of the event's $(u = 2)$ occurrence is equal $\pi_X(2) = 1$. The greatest possible probability of the event "Hans ate 3 eggs for breakfast" results from the probability distribution $p_{X6}(u)$ in Fig. 8f and is equal 0.2. It means that the possibility of the event's $(u = 3)$ occurrence equals $\pi_X(3) = 0.2$.

The smallest possible probability that Hans ate one egg for breakfast $p_X(u = 1) = 0$ corresponds to the probability distributions $p_{X1}(u)$, $p_{X2}(u)$, and $p_{X3}(u)$ in Figs. 8d, e, f. It means that the **necessity of the event's** $(u = 1)$ **occurrence** $\eta(u = 1) = 0$ (this event could not take place). The smallest possible probability of the event "Hans ate two eggs for breakfast" equals $p_X(u = 2) = 0.8$ (according to the information "A" Hans ate at least for 8 of all 10 breakfasts 2 eggs). It results from the probability distributions $p_{X1}(u)$, $p_{X3}(u)$, and $p_{X6}(u)$ in Figs. 8a, c, f. This means that the necessity of the event's $(u = 2)$ occurrence equals $\eta_X(u = 2) = 0.8$. In the case of the event "Hans ate three eggs for breakfast" the smallest possible probability $p_X(u = 3) = 0$ results from the probability distributions $p_{X1}(u)$, $p_{X2}(u)$, $p_{X4}(u)$ in Figs. 8a, b, d. It means that the necessity of this event equals $\eta_X(3) = 0$. All the particular distributions are presented in Fig. 9.

The necessity distribution $\eta_X(u)$ of the event occurrence in the considered simple example with only 3 possible values of the variable u $(u = 1, 2, 3)$ and with only 10 measurements I_k (10 breakfasts) was easy to determine without the use of mathematical formulas. In more complicated problems we may apply

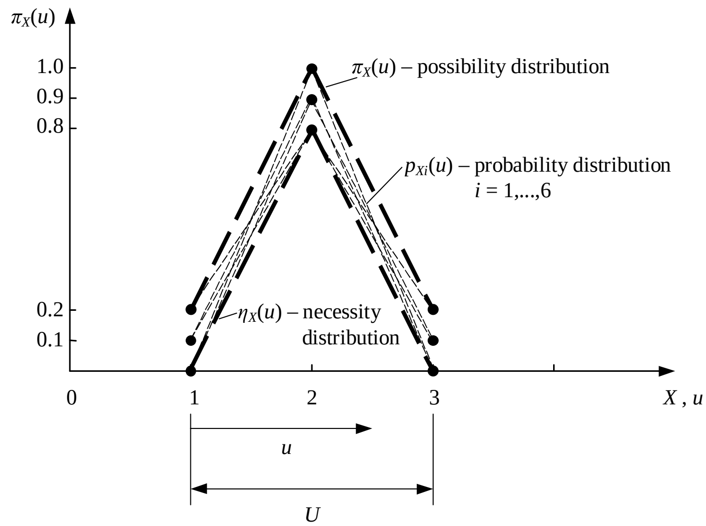

Figure 9. Possibility distribution $\pi_X(u)$, probability distributions $p_{Xi}(u)$, $i = 1 \div 6$, and necessity distribution $\eta_X(u)$ of the event "Hans ate X eggs for breakfast"

formula (2):

$$N(A) = 1 - \Pi(\bar{A})$$
$$N(A) = \inf\{1 - \pi(u) \mid u \notin A\}$$

For A = {1}: N({1}) = 1-1 = 0.
For A = {2}: N({2}) = 1-0.2 = 0.8.
For A = {3}: N({3}) = 1-1 = 0.

The above calculations, carried out with the formula (2) confirm the results of necessity evaluation made on the basis of all possible probability distributions $p_{Xi}(u)$ of the event "Hans ate X eggs for breakfast". Because in the considered problem several (6) different probability distributions are possible, an "average" distribution $p_{Xaver}(u)$ of probability can be determined on the basis of analysis of probability values for particular values of the variable u in Figs. 8a,b,c,d,e,f.

$$\begin{aligned}
p_{Xaver}(u=1) &= \frac{1}{6}\,(0.2+0.1+0.1+0+0+0) = 0.067\\
p_{Xaver}(u=2) &= \frac{1}{6}\,(0.8+0.9+0.8+1.0+0.9+0.8) = 0.866\\
p_{Xaver}(u=3) &= \frac{1}{6}\,(0+0+0.1+0+0.1+0.2) = 0.067. \qquad (14)
\end{aligned}$$

In practical tasks we have to do with much more complicated problems, in which we are not able to determine all possible probability distributions $p_{Xi}(u)$. Therefore, in order to determine the "average" possibility distribution we may use formula (9) which uses for this purpose the possibility distribution $\pi_X(u)$ of the event:

$$p_{Xaver}(u_i) = \sum_{j=i}^{n} \frac{1}{j}\{\pi(u_j) - \pi(u_{j+1})\}.$$

Particular possible values of the variable u are ordered here according to formula (9): $u_1 = 2$, $u_2 = 1$, $u_3 = 3$. Calculations are given below.

$$\begin{aligned}
p_{Xaver}(u_1 = 2) &= \sum_{j=1}^{3} \frac{1}{j}\{\pi_X(u_j) - \pi_X(u_{j+1})\} \\
&= \frac{1}{1}(1 - 0.2) + \frac{1}{2}(0.2 - 0.2) + \frac{1}{3}(0.2 - 0) = 0.866 \\
p_{Xaver}(u_2 = 1) &= \sum_{j=2}^{3} \frac{1}{j}\{\pi_X(u_j) - \pi_X(u_{j+1})\} \\
&= \frac{1}{2}(0.2 - 0.2) + \frac{1}{3}(0.2 - 0) = 0.067 \\
p_{Xaver}(u_3 = 3) &= \sum_{j=3}^{3} \frac{1}{j}\{\pi_X(u_j) - \pi_X(u_{j+1})\} \\
&= \frac{1}{3}(0.2 - 0) = 0.067. \qquad (15)
\end{aligned}$$

It turns out that the "average" probability distributions $p_{Xaver}(u)$ of the event "Hans ate X eggs for breakfast" given by two formulas (14) and (15) are identical ones. Fig. 10 shows the possibility distribution $\pi_X(u)$, the "average" probability distribution $p_{Xaver}(u)$, and the necessity distribution $\eta_X(u)$ of the event.

Using the "average" probability distribution $p_{Xaver}(u)$ seems necessary in more complicated practical problems. However, in the case of a so simple example as the one analyzed in this chapter, where all the feasible probability distributions $p_{Xi}(u)$ are known, the problem can be investigated very precisely. The number of feasible probability distributions $p_{Xi}(u)$ equals 6, see Fig. 8. Analysis of the distributions allows us to discover that the probability value $p_X(u = 1) = 0.2$ occurs in only one out of all 6 distributions: in $p_{X1}(u)$. It means that **probability** $p(p_X(u = 1) = 0.2)$ **of occurrence of the probability value** $p_X(u = 1) = 0.2$ is equal 1/6.

Next, the probability value $p_X(u = 1) = 0.1$ occurs in two out of all 6 probability distributions, in $p_{X2}(u)$ and $p_{X3}(u)$. The probability $p(p_X(u = 1) = 0.1)$ of occurrence of the probability value $p_X(u = 1) = 0.1$ is therefore

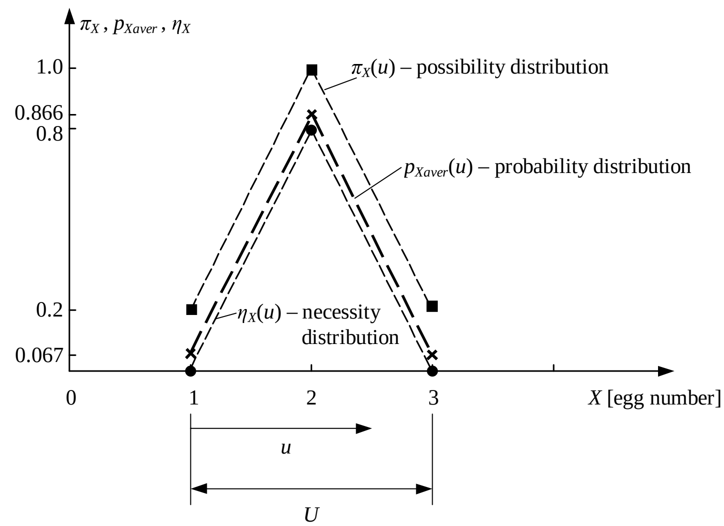

Figure 10. Possibility distribution $\pi_X(u)$, the "average" probability distribution $p_{X\,aver}(u)$, and necessity distribution $\eta_X(u)$ of the event "Hans ate X eggs for breakfast"

equal 2/6. The probability value $p_X(u = 1) = 0$ occurs in three out of all 6 probability distributions, in $p_{X4}(u = 1)$, $p_{X5}(u = 1)$ and in $p_{X6}(u = 1)$. It means that probability $p(p_X(u = 1) = 0)$ of occurrence of the probability value $p_X(u = 1) = 0$ equals 3/6.

In a similar manner we may calculate the probability of occurrence of probability values of Hans eating two eggs: $p(p_X(u{=}2))$, and three eggs: $p(p_X(u{=}3)$. Probability distributions of probability values for particular events $(u = 1)$, $(u = 2)$ and $(u = 3)$ are presented in Fig. 11.

Let us notice in Fig. 11 that probability of occurrence of the maximum probability value $p_X(u = i) = \pi_X(u = i)$, that is, of the probability value, which is equal to the possibility of the event $(u = i)$ is for all values $i = 1, 2, 3$ the smallest one and equal to 1/6, while the probability of the necessity $\eta_X(u = i)$ is for all i the greatest one and equal to 3/6.

When we use the possibility $\pi_X(u)$ and the necessity $\eta_X(u)$ we have to do with a little strange notion of *probability of occurrence of probability value of the event* $(X = u)$ which causes that the possibility distribution is no longer 2-dimensional but becomes a 3-dimensional distribution. For the analyzed "Hans/eggs"-problem this distribution is shown in Fig. 12.

In the case of fuzzy sets there are known and used the type 2-fuzzy sets (Zimmermann, 1996), for which the membership values in a set are the type 1-fuzzy sets. Similarly, in the investigated "Hans/eggs"-problem we have to

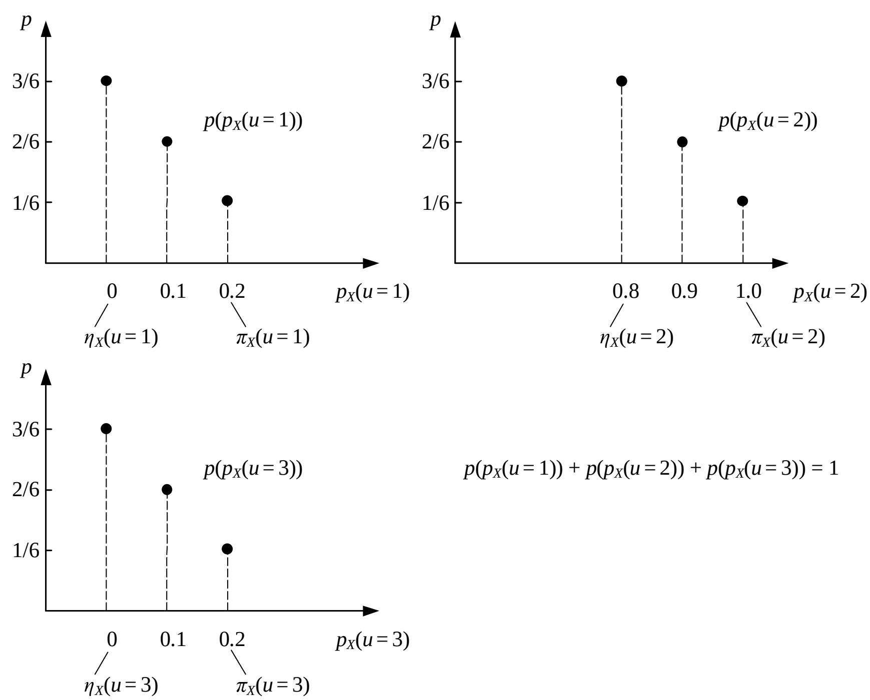

Figure 11. Possibility distributions $p(p_X(u = i))$, $i = 1, 2, 3$ of occurrence of probability value $p_X(u = i) = 0.1$ or 0.2 or 0.3 of the event "Hans ate X eggs for breakfast" resulting from 6 possible probability distributions $p_{Xi}(u)$ presented in Fig. 8

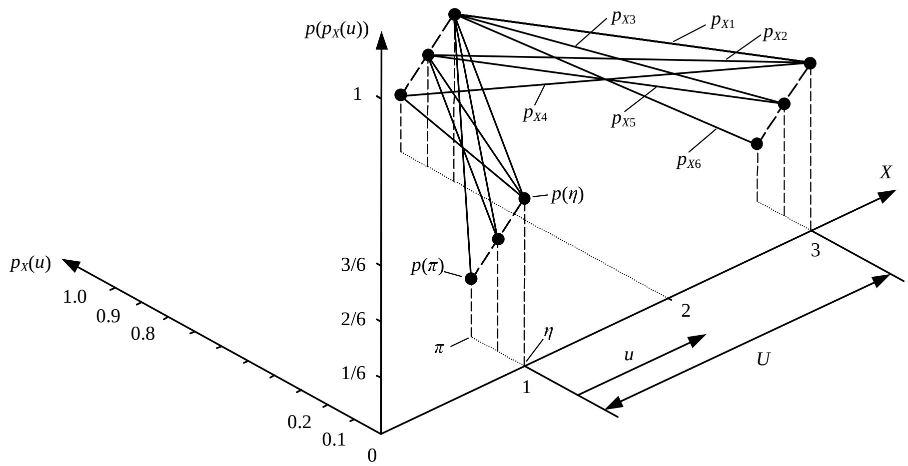

Figure 12. Three-dimensional probability distribution of the event $(X = u)$: "Hans ate X eggs for breakfast", i.e. the probability distribution $p(p_X(u))$ of occurrence of the probability value $p_X(u)$ of the event $(X = u)$

do with type 2-probability in which the probability value $p_X(u)$ of the event $(X = u)$ is the type 1-probability. Because determining type 2-probability distributions is rather difficult, in practical applications we have to use the simplified, "average" type 1-probability distributions $p_{Xaver}(u)$.

5. Conclusions

If evidence information about the problem is given in the form of nested focal elements $E_1 \subseteq E_2 \subseteq \ldots E_p$ then the event $(X = u)$ cannot be directly described with only one probability distribution $p_X(u)$ because the uncertainty of the evidence information (uncertainty of measurements) causes that not one but many probability distributions $p_{Xi}(u)$ are possible. Then, for each possible event $(X = u)$ its possibility $\pi_X(u)$ can be determined. This possibility is the greatest possible probability of the event $(X = u)$ resulting from only one of many possible probability distributions $p_{Xi}(u)$. Because probability that such specific distribution $p_{Xi}(u)$ really took place in the considered problem equals $1/m$, where m is number of all possible distributions $p_{Xi}(u)$ and is usually very large, the value of $1/m$ is usually very small. In the simple "Hans/eggs"-example this probability equals 1/6. But when the variable X is continuous, the number of all possible distributions is infinitely large and the probability of occurrence of one specific distribution is infinitely small. It means that **the possibility** $\pi_X(u)$ **refers to a very little probable case, to the extreme case**. Therefore **the informative value of possibility** $\pi_X(u)$ **of the event's** $(X = u)$ **occurrence**

is small. The possibility should be used as an additional information but not as the main information about the problem under consideration.

A similar conclusion refers to the necessity $\eta_X(u)$ of the event occurrence. The necessity informs us about the minimal but certain probability value of the event occurrence. When we know values of the possibility $\pi_X(u)$ and of the necessity $\eta_X(u)$ then we can be certain that the event $(X = u)$ occurs with probability at least equal to the necessity but not exceeding the possibility. The use in the solution of real problems of the measures of possibility and necessity of an event occurrence corresponds, respectively, to the extremely optimistic and to the extremely pessimistic approach to the problem. Both approaches are extreme and mostly of small practical value.

The possibility that Hans will for all successive 1000 breakfasts eat always two eggs equals 1. But probability of such a sequence of events is small and equals only 1/6 (6 probability distributions in the problem are possible). Basing the problem solution on probability is more practical. The probability that Hans will eat two eggs for breakfast equals 0.866 (Fig. 10) and is much higher than 1/6. Using it, we can hope that Hans for 866 out of 1000 breakfasts will eat two eggs. It is much more certain and valuable information than possibility.

The above conclusions have a very great meaning for identification of membership functions of fuzzy sets and for fuzzy arithmetics. If fuzzy numbers which represent real variables like, e.g., approximate income and costs of a firm (e.g. about 7 million Euro and about 5 million Euro) were identified as possibilistic fuzzy numbers then the calculated value of *profit = income - costs*, being fuzzy number about 2 million Euro, will also have the character of possibility distribution $\pi_X(\mathit{profit})$ and will inform us about some extreme events, whose occurrence is possible but very little probable. The use of the possibilistic fuzzy numbers results in calculation paradoxes and produces too fuzzified results. The present author showed in Piegat (2003) that in subtraction of fuzzy numbers an insensitivity of the subtraction result on numbers successively subtracted from the minuend occurs. Therefore the author proposes in fuzzy arithmetic the application of probabilistic fuzzy numbers instead of the possibilistic ones and instead of the possibilistic extension principle of Zadeh or Klir (Klir, 1997), the use of the cardinality extension principle (Piegat, 2003). This principle gives results of greater informative value, which are more suitable for practical applications. If a fuzzy number is of possibilistic character, then it is recommended (using formula (9)) to calculate the probability distribution corresponding to the possibility distribution and next to normalize it to interval $[0, 1]$. In this way we get a probabilistic fuzzy number for which the cardinality extension principle can be used in calculations. The transformation of the possibilistic fuzzy number *about 2* (eggs), referring to the "Hans/eggs"- example, into a probabilistic fuzzy number *about 2* (eggs) is shown in Fig. 13.

Arithmetic operations on probabilistic fuzzy numbers with the use of cardinality extension principle give results, which radically differ from the results achieved with possibilistic fuzzy numbers, calculated with the possibilistic ex-

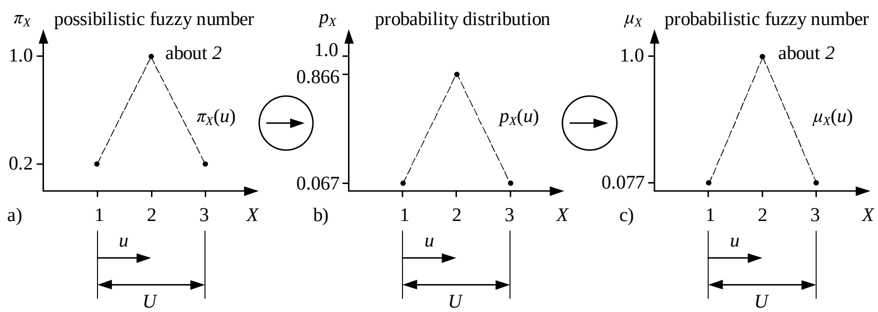

Figure 13. Example of transformation of the possibilistic fuzzy number *about 2* (eggs) – (a) into the probabilistic fuzzy number *about 2* (eggs) – (c)

tension principle of Zadeh or Klir (Klir, 1997). These results are also more suitable for practice.

References

Borgelt, C. and Kruse, R. (2003) Learning possibilistic graphical models from data. *IEEE Transactions on Fuzzy Systems* **11** (2).

Driankov, D., Hellendoorn, H. and Reinfrank, M. (1993) *An Introduction to Fuzzy Control.* Springer-Verlag Berlin, Heidelberg.

Dubois, D. and Prade, H. (1988) *Possibility Theory. An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York and London.

Dubois, D. and Prade, H. (1993) Fuzzy sets and probability: misunderstandings, bridges and gaps. *Proceedings of the Second IEEE Conference on Fuzzy Systems*, 1059-1068.

Dubois, D. and Prade, H. (1994) Possibility Theory, Belief Revision and Nonmonotonic Logic. *Lecture Notes in Computer Science* **847**, 51-74.

Dubois, D. and Prade, H. (2002) Quantitative possibility theory and its probabilistic connections. In: P. Grzegorzewski, O. Hryniewicz, M.A. Gil, eds., *Soft Methods in Probability, Statistics and Data Analysis.* Physica Verlag, Heidelberg-Germany, Advances in Soft Computing, 3-26.

Dubois, D., Foulloy, L., Mauris, G. and Prade, H. (2004) Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing* **10** (4), 273-297.

Kacprzyk, J. (1986) *Zbiory rozmyte w analizie systemowej (Fuzzy sets in system analysis).* Państwowe Wydawnictwo Naukowe, Warszawa.

Klir, G.J. and Folger, T.A. (1988) *Fuzzy Sets, Uncertainty and Information.* Prentice-Hall International Inc., Englewood Cliffs, New Jersey.

Klir, G.J. (1997) Fuzzy arithmetic with requisite constraints. *Fuzzy Sets and*

Systems **91**, 165-175.

PIEGAT, A. (2003) Informative value of the possibilistic extension principle. *Proceedings of* 10^{th} *International Multi-Conference on Advanced Computer Systems ACS 2003*, Międzyzdroje, Poland, October 22-24, CD (also to be published by Kluwer).

RUSPINI, E.H. (1990) Similarity based models for possibilistic logics. *Proceedings of the 3rd International Conference on Information Processing and Management Uncertainty Knowledge Based Systems*, Granada, Spain, 56-58.

SPOHN, W. (1990) A general nonprobabilistic theory of inductive reasoning. In: Schachter R.D., Levitt T.S., Kanal N. and Lemmer J.F., eds., *Uncertainty in Artificial Intelligence* **4**, Amsterdam, Netherlands, 149-159.

SPOTT, M. (1997) A theory of possibility distributions. *Fuzzy Sets and Systems* **102**, 135-155.

ZADEH, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* **1**, 3-28.

ZIMMERMANN, H.-J. (1996) *Fuzzy Set Theory and its Applications.* Kluwer Academic Publishers, Boston/Dordrecht/London.