

Fuzzy Knowledge Representation Using Probability Measures of Fuzzy Events

Anna Walaszek-Babiszewska
Opole University of Technology
Poland

1. Introduction

The concepts of *Soft Computing* introduced by Lotfi A. Zadeh in 1991 has integrated different methodologies and approaches, as: fuzzy set theory, fuzzy logic, approximate reasoning, linguistic expression of knowledge, probabilistic reasoning, and others for solving problems of complex systems in the way similar to human perception, recognition and solving problem methods. Linguistic fuzzy modelling gives the formal, mathematical instruments for expressing human knowledge described in natural language. Probability of fuzzy meanings of linguistic variables determines a frequency of the occurrence the imprecisely expressed events.

This work presents the methods of applications linguistic modelling and probability measures of fuzzy events for creating models compatible to the features of real systems, and more flexible than traditional rule-based models derived from linguistic knowledge.

In Section 2. we remind the notions of a linguistic variable and a probability of fuzzy events, formulated by Zadeh, which have become fundamental for the development of fuzzy systems. We define probability distributions of a linguistic variable and a linguistic vector as well as a mean fuzzy value (a mean fuzzy set) of the linguistic variable. The conditional probability of fuzzy events will be the base for the inference procedure.

Section 3. shows an exemplary probabilistic modelling for the characteristics representing features of particles in a certain population, formulated in fuzzy categories.

The created knowledge representation states a collection of weighted rules (Section 4.) Weights of rules represent probabilities of fuzzy events of input and output system variables. Construction of fuzzy models is presented for different stochastic systems. The weights are involved in the inference procedures.

2. Linguistic and probabilistic modelling - basic definitions and concepts

2.1 Linguistic variable

Linguistic fuzzy modelling with a linguistic knowledge representation gives us the formal, mathematical way for expressing the human linguistic perception of real world. A linguistic variable is the main notion in such modelling. The variable whose values are words, can be defined by the quadruplet $\langle X, L(X), U, M_x \rangle$, where X means the name of the variable, $L(X)$ is a set of linguistic values (words) which X takes, U is an universe of discourse, and M is a

semantic function, that assigns the fuzzy meaning (fuzzy subsets $A_i, i=1, \dots, l$ in U) to each linguistic value from $L(X)$ (Zadeh, 1975).

As an example let us consider a linguistic variable $X = \text{diameter of particles}$, which takes the linguistic values $L(X) = \{\text{fine, middle size, coarse}\}$. A semantic function M associates fuzzy meanings (fuzzy sets) $A_i, i=1, 2, 3$ to each linguistic value. Fuzzy sets are given by membership functions $\mu_A(u)$ determined in the universe of discourse $U \subset R$ (an interval of real numbers), as follows:

$$\mu_{A_1}(u) = \begin{cases} 1, & u \in [0, 0.2) \\ \frac{0.35 - u}{0.15}, & u \in [0.2, 0.35] \end{cases} \quad (1)$$

$$\mu_{A_2}(u) = \begin{cases} \frac{u - 0.2}{0.15}, & u \in [0.2, 0.35) \\ 1, & u \in [0.35, 0.65] \\ \frac{0.8 - u}{0.15}, & u \in (0.65, 0.8] \end{cases} \quad (2)$$

$$\mu_{A_3}(u) = \begin{cases} \frac{0.65 - u}{0.15}, & u \in [0.65, 0.8) \\ 1, & u \in [0.8, 1] \end{cases} \quad (3)$$

where A_1 : fine, A_2 : middle size, A_3 : coarse.

2.2 Probability of fuzzy events and distributions of linguistic variables

A fuzzy event A , in compliance with Zadeh's definition, is a fuzzy subset $A = \{(\mu_A(u), u)\}$ in the elementary events domain (U, Ω) , with the membership function $\mu_A(u) \in [0, 1]$ measurable in Borel sense (Zadeh, 1968). In another notation, a fuzzy event can be expressed as $A = \sum_{u \in U} (\mu_A(u) / u)$, where the sum sign is considered as set character, not arithmetical.

Let p be the probability function, which assigns to each Borel set in the domain $U = \{u_1, \dots, u_N\}$ the real number $p \in [0, 1]$. A probability $P(A)$ of a fuzzy event A is defined in the way:

$$P(A) = \sum_{u \in U} p(u) \mu_A(u) \quad (4)$$

where $p(u) \in [0, 1]$ is a probability function in (U, Ω) .

If the domain U is infinite, and $f(u)$ is a probability density function in (U, Ω) , then a probability $P(A)$ of a fuzzy event A can be calculated by the following integral:

$$P(A) = \int_{u \in U} \mu_A(u) f(u) du \quad (5)$$

Probability values of fuzzy events, calculated according to (4) and (5) are real numbers, $P(A) \in [0, 1]$.

Let fuzzy subsets $A_i, i = 1, \dots, I$ of the linguistic variable X are defined by their membership functions $\mu_{A_i}(u) \in [0, 1]$ in such way, that for every $u \in U$, the relationship $\sum_{i=1}^I \mu_{A_i}(u) = 1$ is fulfilling. Then, the set of probabilities $\{P(A_i)\}, i = 1, \dots, I$ calculated for fuzzy subsets $A_i, i = 1, \dots, I$ according to (4) or (5) and fulfilling the relation

$$\sum_{i=1}^I P(A_i) = 1 \tag{6}$$

states a probability distribution of the linguistic variable X .

A mean fuzzy value of the linguistic random variable X , signed as \bar{A} , in a probability distribution $P(A_i), i = 1, \dots, I$ given above, is a fuzzy set with the membership function defined as follows:

$$\mu_{\bar{A}}(u) = P(A_1)\mu_{A_1}(u) + \dots + P(A_I)\mu_{A_I}(u), \forall u \in U \tag{7}$$

The mean fuzzy value is the convex combination fuzzy set (Kacprzyk, 1986). It will be used in the aggregation procedure (Section 4).

For the linguistic variable whose fuzzy meanings of linguistic values are given by (1) - (3), assume that probability density function is constant over the domain $U = [0, 1]$, $f(u) = 1$. Probability of fuzzy events A_1 : *fine*, A_2 : *middle size*, A_3 : *coarse* can be calculated by the integrals

$$P(A_1) = \int_0^{0.35} \mu_{A_1}(u) f(u) du = 0.275, \tag{8}$$

$$P(A_2) = \int_{0.2}^{0.8} \mu_{A_2}(u) f(u) du = 0.450, \tag{9}$$

$$P(A_3) = \int_{0.65}^1 \mu_{A_3}(u) f(u) du = 0.275. \tag{10}$$

The calculated probability values fulfil (6), so the mean fuzzy value of events A_1, A_2, A_3 can be calculated according to the dependence (7).

Let us consider, as it is usually assumed in industrial practice, that for disjoint intervals $\Delta u_m = a_m, m = 1, \dots, M$ the empirical probability is constant and equal to the quotient

$$p(\Delta u_m) = P(u \in \Delta u_m) = \frac{n_m}{n} \tag{11}$$

where n_m is the number of particles whose a feature x takes the values from the proper intervals, and n is the total number of particles, $n = \sum_{m=1, \dots, M} n_m$. Fuzzy sets $A_i, i=1, \dots, I$ representing linguistic values of a linguistic variable X can be determined on the disjoint intervals a_m as follows

$$A_i = \sum_{m=1, \dots, M} \mu_{A_i}(a_m) / a_m \quad (12)$$

where the membership functions fulfill the condition $\sum_{i=1, \dots, I} \mu_{A_i}(a_m) = 1, \forall a_m \in U$.

Probabilities of the events (12) can be expressed by the relationship

$$P(A_i) = \sum_{m=1, \dots, M} \mu_{A_i}(a_m) p_m \quad (13)$$

where $p_m = P(u \in a_m)$. If the set of probabilities $\{P(A_i)\}$ of fuzzy events, calculated according to (12) and (13) is fulfilling the relationship $\sum_{i=1, \dots, I} P(A_i) = 1$ then it states a probability distribution of the linguistic variable X .

Let us consider two linguistic variables: $\langle X, L(X), U, M_x \rangle$ and $\langle Y, L(Y), V, M_y \rangle$, where X means the name of the input (reason) variable, Y is the name of the output (result) variable of a SISO system, $L(X)$ and $L(Y)$ are sets of linguistic values, $U \subset R$ and $V \subset R$ are universes of discourse of particular variables, M_x, M_y are semantic functions, that assign fuzzy meaning (fuzzy subsets $A_i, i=1, \dots, I$ in $U \subset R$ and $B_j, j=1, \dots, J$ in $V \subset R$) to each linguistic value from $L(X)$ and $L(Y)$, respectively.

Fuzzy sets $A_i, i=1, \dots, I$ and $B_j, j=1, \dots, J$ state the numeric descriptions of particular linguistic values of variables X and Y , respectively. Let fuzzy sets are defined by membership functions in the way

$$\mu_{A_i}(u) \in [0, 1], i=1, \dots, I \text{ and } \sum_{i=1}^I \mu_{A_i}(u) = 1, \forall u \in U; \quad (14)$$

$$\mu_{B_j}(v) \in [0, 1], j=1, \dots, J \text{ and } \sum_{j=1}^J \mu_{B_j}(v) = 1, \forall v \in V. \quad (15)$$

A new term set of a linguistic vector (X, Y) can be created in the space $L(X) \times L(Y)$ by the numeric description $\mu_{A_i \times B_j}(u, v)$ in the universe $U \times V$. The membership function $\mu_{A_i \times B_j}(u, v)$ is a Borel function, which can be defined as a t-norm

$$\mu_{A_i \times B_j}(u, v) = T(\mu_{A_i}(u), \mu_{B_j}(v)) \quad (16)$$

e.g., as a product t-norm $\mu_{A_i \times B_j}(u, v) = \mu_{A_i}(u) \mu_{B_j}(v)$.

A set of probability values $\{P(A_i \times B_j)\}$, of fuzzy events (fuzzy relations) $A_i \times B_j, i=1, \dots, I, j=1, \dots, J$ can be calculated using the basic definition (4), as follows

$$P(A_i \times B_j) = \sum_{(u,v) \in U \times V} p(u,v) \mu_{A_i \times B_j}(u,v) \tag{17}$$

where $p(u,v) \in [0,1]$ is a joint probability function.

If the set of probabilities of fuzzy events $\{P(A_i \times B_j)\}$ fulfils the relationship

$$\sum_{i=1}^I \sum_{j=1}^J P(A_i \times B_j) = 1 \tag{18}$$

then, it states a joint probability distribution $P(X,Y)$ of the linguistic vector variables (X,Y) . The marginal probability distributions $P(X) = \{P(A_i)\}, i = 1, \dots, I$, and $P(Y) = \{P(B_j)\}, j=1, \dots, J$ can be calculated for particular linguistic variables, as follows:

$$P(A_i) = \sum_{j=1}^J P(A_i \times B_j) \tag{19}$$

$$P(B_j) = \sum_{i=1}^I P(A_i \times B_j) \tag{20}$$

The marginal distributions $P(X)$ and $P(Y)$ defined above are normalized:

$$\sum_{i=1}^I P(A_i) = 1, \quad \sum_{j=1}^J P(B_j) = 1 \tag{21}$$

Conditional probability distributions of particular linguistic variables can be derived from the joint probability distribution $P(X,Y)$ given by (17) and a marginal probability distributions (19) or (20). The conditional probability distribution $P(Y/X)$ of the linguistic variable Y , under the condition that X takes fuzzy values $A_i, i=1, \dots, I$, is a set of probability values $P(Y/X) = \{P(B_j / A_i)\}, j=1, \dots, J; i=const$, calculated as follows:

$$P(B_j / A_i) = P(A_i \times B_j) / P(A_i), j=1, \dots, J; i=const. \tag{22}$$

Taking into account the normalization of probability distributions (see (18) and (21)), the total probability of the result B_j can be calculated, similarly to Bayes' formula, and assuming, that the conditional probabilities $P(B_j / A_i), i=1, \dots, I$, calculated under the condition of the reasons A_i are known (Walaszek-Babiszewska, 2008):

$$P(B_j) = \sum_{i=1}^I P(B_j / A_i) P(A_i). \tag{23}$$

Let us note, that fuzzy sets $A_i, i=1, \dots, I$ are not disjoint in U .

3. Exemplary linguistic and probabilistic modelling

3.1 Particle characteristics as results of measurements

In chemical and biochemical research, in many industrial processes such as mineral preparation processes or in numerous food processes, the material to be prepared consists of a population of different types of particles.

There are basic characteristics of material utilised by process engineers and automation engineers:

- a characteristic of the size composition presenting portions of particles belonging to different size fractions,
- a densimetric characteristic presenting portions of particles belonging to different density fractions,
- a characteristic of tested chemical components.

The two first characteristics are often considered as an empirical probability distribution of a two-dimensional random variable: volume x and density y of particles

$$p_{ij}(x, y) = P(x \in a_i, y \in b_j) = N_{ij} / N, \quad i = 1, 2, \dots, I; \quad j = 1, 2, \dots, J \quad (24)$$

where $a_i, i=1, \dots, I$ are disjoint intervals of the particle size (size classes) in a domain X , and $b_j, j=1, \dots, J$ are disjoint intervals of the particle density (density fractions) in a domain Y , N_{ij} is a number of that particles in the parent population, whose volume belongs to i -th interval a_i and the density belongs to j -th interval b_j , N is a total number of particles in the population,

$$\text{and } N = \sum_{i=1}^I \sum_{j=1}^J N_{ij}.$$

In engineering practice a different measure of the probability is being more often applying, the quotient of the respective masses:

$$\pi_{ij}(x, y) = M_{ij} / M \quad i=1, 2, \dots, I; \quad j=1, 2, \dots, J; \quad (25)$$

where M_{ij} is a mass of that particles in the parent population, whose volume belongs to i -th interval a_i and the density belongs to j -th interval b_j , and $M_{ij} = x_{m,i} y_{m,j} N_{ij}$, where $x_{m,i}$, $y_{m,j}$ are mean values of particle volume and density in the intervals a_i and b_j , respectively; M is a total mass of the population, $M = x_m y_m N$, where x_m , y_m are mean values of volume

and density of particles in the population, and $M = \sum_{i=1}^I \sum_{j=1}^J M_{ij}$.

There is a relation between two expressions of empirical probabilities (Walaszek-Babiszewska, 2004):

$$\pi_{ij} = a_{ij} p_{ij}, \quad i=1, 2, \dots, I; \quad j=1, 2, \dots, J; \quad (26)$$

where:

$$a_{ij} = \frac{x_{m,i} y_{m,j}}{x_m y_m}. \quad (27)$$

Table 1 presents the values of an empirical joint probability distribution $p(x_i, y_j) = p_{ij}$, calculated on the base of measurements (Walaszek-Babiszewska, 2004). Each of the ranges of

volume and density of particles has been divided into 4 intervals. The smallest value of indexes i, j concerns to the smallest value of density and volume:

$$x_1 < x_2 < x_3 < x_4; \quad y_1 < y_2 < y_3 < y_4.$$

The marginal probabilities $p_i(x)$ and $p_j(y)$ are also given in Table 1.

Density fraction number	Probability $p_{ij}(x, y)$				Marginal probability $p_j(y)$
	Size class number				
j	$i=1$	$i=2$	$i=3$	$i=4$	
1	0.6290	0.0304	0.0074	0.0020	0.6688
2	0.1338	0.0044	0.0011	0.0003	0.1396
3	0.0534	0.0050	0.0010	0.0002	0.0596
4	0.1198	0.0092	0.0025	0.0005	0.1320
Marginal probability $p_i(x)$	0.9360	0.0490	0.0120	0.0030	$\sum_i p_i =$ $\sum_j p_{j,j} = 1$

Table 1. The empirical joint probability distribution $p_{ij}(x,y)$ and marginal probability distributions $p_j(y), p_i(x)$ of particle features

3.2 Linguistic characteristics of particles

Perception-based information of human experts, expressing in natural language a quantity-quality characteristic of particles features, can be verified on the base of linguistic and probabilistic modelling presented above.

Let us assume two linguistic variables considered above:

x_{name} : 'volume of particles' and the set of linguistic values $L(X)=\{small(A_1), middle(A_2), large(A_3)\}$ with the membership functions determined over the disjoint intervals $a_i, i=1, \dots, I$ as follows:

$$A_1 = 1/a_1 + 0.2/a_2; \quad A_2 = 0.8/a_2 + 0.8/a_3; \quad A_3 = 0.2/a_3 + 1/a_4$$

y_{name} : 'density of particles' and the set of linguistic values $L(Y)=\{light(B_1), middle(B_2), heavy(B_3)\}$ with the membership functions determined over the disjoint intervals $b_j, j=1, \dots, J$, as follows:

$$B_1 = 1/b_1 + 0.5/b_2; \quad B_2 = 0.5/b_2 + 0.5/b_3; \quad B_3 = 0.5/b_3 + 1/b_4$$

For these two linguistic variables and the empirical joint probability distribution given in Table 1. we can calculate the probability of the simultaneous events, e.g.

$$P\{ (x \text{ is small}) \text{ and } (y \text{ is light})\}$$

using (16) i (17) as follows:

$$P(A_1 \times B_1) = p_{11}(x, y)\mu_{A_1}(a_1)\mu_{B_1}(b_1) + p_{12}(x, y)\mu_{A_1}(a_1)\mu_{B_1}(b_2) + p_{21}(x, y)\mu_{A_1}(a_2)\mu_{B_1}(b_1) + p_{22}(x, y)\mu_{A_1}(a_2)\mu_{B_1}(b_2)$$

$$P(A_1 \times B_1) = 0.6290 \cdot 1 \cdot 1 + 0.1338 \cdot 1 \cdot 0.5 + 0.0304 \cdot 0.2 \cdot 1 + 0.0044 \cdot 0.2 \cdot 0.5 = 0.7024$$

In the similar way we can compute the values of probabilities $P(A_i \times B_j)$, $i, j=1, 2, 3$ (Table 2).

	A_1	A_2	A_3	$P(B_j)$
B_1	0.7024	0.0324	0.0037	0.7385
B_2	0.0946	0.0046	0.0005	0.0997
B_3	0.1488	0.0118	0.0012	0.1618
$P(A_i)$	0.9458	0.0488	0.0054	$\sum P(A_i) =$ $\sum P(B_j) = 1$

Table 2. The probability distributions of linguistic variables (x, y) representing probability of fuzzy events $P(A_i \times B_j)$ and marginal probabilities

Probability distributions of linguistic variables (Table 2.) could be used for the validation of experts' opinion:

'The contents of the light and small particles is very high'; $P(A_1 \times B_1) = 0.7024$;

'The contents of the light fraction is high'; $P(B_1) = 0.7385$;

'The contents of large particles is low'; $P(A_3) = 0.0054$.

The values of probability of fuzzy events calculated according to (16) and (17) depend on a choice of a t-norm. The problem is important for creating the inference procedure in knowledge-based systems.

3.3 Quality parameters of particles as a mean value of a fuzzy event

Characteristics of tested chemical components in population of particles are usually called the quality characteristics. Suppose the quality parameter $\beta_{ij} = \beta(x_i, y_j)$ in every elementary fraction of particles, where x_i, y_j are mean values of particle volume and density in the intervals a_i and b_j , respectively. The mean value of a tested substance in the population of particles whose features are determined by a fuzzy event e.g. "C: small and light particles" can be calculated by using the notion of a mean value of fuzzy event, as follows (Walaszek-Babiszewska, 2004):

$$\beta_C = \frac{\sum_{i=1}^I \sum_{j=1}^J \beta_{ij} p_{ij} \mu_C(x_i, y_j)}{\sum_{i=1}^I \sum_{j=1}^J p_{ij} \mu_C(x_i, y_j)} \quad (28)$$

4. Knowledge representation

4.1 General form of a fuzzy model

A fuzzy model represented a MISO system, consisting of the collection of fuzzy rules in a form 'IF x is A THEN y is B ' is considered (Yager and Filev, 1995). The propositions x is A in antecedents and y is B in consequents of rules are based on the partition of the input-output space, given by experts. The exemplary i -th file rule includes J elementary rules, in a form:

$$\begin{aligned}
 &R_i: w_i (IF x_1 \text{ is } A_{1,i} \text{ AND } x_2 \text{ is } A_{2,i} \dots \text{ AND } x_p \text{ is } A_{p,i} \\
 &\text{ THEN } y \text{ is } B_1 (w_{1/i}) \\
 &\text{-----} \\
 &\text{ ALSO } y \text{ is } B_j (w_{j/i}) \\
 &\text{-----} \\
 &\text{ ALSO } y \text{ is } B_J (w_{J/i})
 \end{aligned}
 \tag{29}$$

where $x^T = (x_1, \dots, x_p)$ is a linguistic vector of system inputs (antecedent variables), y is a linguistic output variable. The linguistic values (term sets) $L(x_1), \dots, L(x_p), L(y)$ of the input and output variables are predefined by process experts. Fuzzy sets $A_{1,i}, \dots, A_{p,i}$, $i=1, \dots, I$ represent linguistic values of the input vector, and are defined by membership functions $\mu_{A_{1,i}}(u_1), \dots, \mu_{A_{p,i}}(u_p)$ in the domains U_1, \dots, U_p , $U_i \subset R, i=1, \dots, p$. The linguistic output (consequent) variable y has the family of fuzzy subsets B_j with membership functions $\mu_{B_j}(v), j=1, \dots, J$ in the numeric space $V \subset R$.

The rule weights w_i and $w_{j/i}, i=1, \dots, I, j=1, \dots, J$ represent probabilities of fuzzy events occurring in the antecedents and consequence of the model $w_i = P(A_i), w_{j/i} = P(B_j / A_i)$ (Walaszek-Babiszewska, 2007). The weight w_i of a file rule represents a joint probability of fuzzy events $A_{1,i} \times \dots \times A_{p,i}$ in the antecedent domain, calculated according to (16) and (17):

$$P(A_i) = \sum_{u \in U} p(u_1, \dots, u_p) T(\mu_{A_{1,i}}(u_1), \dots, \mu_{A_{p,i}}(u_p)) \tag{30}$$

where T means a t-norm, membership functions $\mu_{A_{1,i}}(u_1), \dots, \mu_{A_{p,i}}(u_p)$ are defined in a such way, that for every numeric value $u^{T*} = (u_1^*, \dots, u_p^*) \in U$ the relationship $\sum_{i=1,2,\dots,I} \mu_{A_i}(u^*) = 1$ is fulfilled, and $p(u_1, \dots, u_p)$ is a probability distribution which assigns to each Borel set in U a real number $p \in [0,1]$.

Elementary rule weights $w_{j/i}, i=const, j=1, \dots, J$ state the conditional probabilities of the events (y is B_j) of the consequent variable, under the condition of the input variable (x is A_i). It can be calculated from Bayesian formula (see (22)), as follows

$$w_{j/i} = P(B_j / A_i) = \frac{P(A_i \times B_j)}{P(A_i)}, j=1, \dots, J, i=const \tag{31}$$

where the probability of fuzzy events (fuzzy relations) $A_i \times B_j, i=1, \dots, I, j=1, \dots, J$, determined by the formula

$$P(A_i \times B_j) = \sum_{(u,v) \in U \times V} p(u,v) T(\mu_{A_i}(u), \mu_{B_j}(v)) = w_{ij} \tag{32}$$

states the joint probability distribution of linguistic input-output vector $P(X, Y) = \{P(A_i \times B_j)\}$, $i=1, \dots, I$, $j=1, \dots, J$. The joint probability distribution $p(u, v) \in [0, 1]$ determined in the input-output universe $U \times V \subset R^{p+1}$ is understood in the sense of the probability theory.

The weights w_i , w_{ij} and $w_{j/i}$ $i=1, \dots, I$, $j=1, \dots, J$ of the model can be estimate by using a set of input-output measurements $\{(u^m, v^m)\}$, $m=1, \dots, M$ and as probability distributions should fulfil the relationships

$$\sum_{i=1}^I w_i = 1, \quad \sum_{i=1}^I \sum_{j=1}^J w_{ij} = 1, \quad \sum_{j=1}^J w_{j/i} = 1, \quad i = const. \quad (33)$$

4.2 Inference and aggregation procedure

The approximate reasoning is based on a fuzzy logic and fuzzy sets theory (Zadeh, 1979). The generalized *modus ponens* permits to deduce an imprecise conclusion from imprecise premises. A great number of works in the literature dealt with fuzzy reasoning, e.g. (Pedrycz, 1984), (Yager & Filev, 1994), (Hellendoorn & Driankov, 1997).

When the proposition x is A_i^* is given, then from the ij -th elementary rule of the model (29), a proposition y is $B_{j/i}^*$ can be computed. The membership function $\mu_{B_{j/i}^*}(v)$ of the inferred fuzzy output is given by the formula:

$$\mu_{B_{j/i}^*}(v) = \sup_{u \in U} T(\mu_{A_i^*}(u), \mu_{R_{ij}}(u, v)) \quad (34)$$

where T means a t-norm, R_{ij} is a fuzzy relation determined in the input-output space $U \times V$ with the membership function $\mu_{R_{ij}}(u, v)$ expressed as an implication operator or as a t-norm (*min* or *product*) derived from membership functions $\mu_{A_i}(u)$ and $\mu_{B_{j/i}}(v)$. Inferred fuzzy set $B_{j/i}^*$ depends on a t-norm as well as the chosen type of the fuzzy relation R_{ij} .

Let us check the inferring procedure from the model (29), taking into account the rule weights representing probabilities of fuzzy events defined above (according to (Walaszek-Babiszewska, 2007a and 2008). Assuming a crisp value (singleton) of input variables $u^* = (u_1^*, \dots, u_p^*)$ with the degree of fitting u^* to the input fuzzy set A_i , calculated by

$$\mu_{A_i}(u^*) = T(\mu_{A_{1,i}}(u_1^*), \dots, \mu_{A_{p,i}}(u_p^*)) = \tau_i \quad (35)$$

the output fuzzy set $B_{j/i}^*$ can be found as follows:

$$\mu_{B_{j/i}^*}(v) = T(\mu_{A_i}(u^*), \mu_{B_{j/i}}(v)) \quad (36)$$

where the t-norm determines the relation R_{ij} . Using the product t-norm (according to Larsen’s rule) in (36), we have the output fuzzy set $B_{j/i}^*$ inferred from ij -th elementary rule, determined by a membership function

$$\mu_{B_{j/i}^*}(v) = \tau_i \mu_{B_{j/i}}(v) \tag{37}$$

Fuzzy outputs $B_{j/i}^*$ computed from elementary rules $j=1, \dots, J$, at the same value of the antecedent ($i=const$), can be aggregated by using weights $w_{j/i}$:

$$\mu_{B_i^*}(v) = \tau_i \sum_{j=1}^J w_{j/i} \mu_{B_{j/i}}(v) \tag{38}$$

The fuzzy set B_i^* derived in such way is a fuzzy conditional mean value (see (7) in paragraph 2.2) of the conclusion (37), calculated under the condition (x is A_i^*).

If the crisp value of input variables $u^* = (u_1^*, \dots, u_p^*)$ belongs also to another input fuzzy sets A_i , and $\tau_i \neq 0, i=1, \dots, I$ then fuzzy outputs of the file rules $B_i^*, i=1, \dots, I$ can be aggregated using weights w_i of all switched rules. Then the aggregated fuzzy set B^* states a total mean fuzzy value of the conclusion (37), with the membership function calculated according to:

$$\mu_{B^*}(v) = \sum_{i=1}^I w_i \tau_i \sum_{j=1}^J w_{j/i} \mu_{B_{j/i}}(v) \tag{39}$$

or in the way:

$$\mu_{B^*}(v) = \frac{\sum_{i=1}^I w_i \tau_i \sum_{j=1}^J w_{j/i} \mu_{B_{j/i}}(v)}{\sum_{i=1}^I w_i} \tag{40}$$

The relationships (33) have been taken into account in formulas (38) and (39).

4.3 Knowledge representation of stochastic systems

Stochastic systems are often described by the ordered pair (x,y) of input and output variables:

$$\{(x(t, \omega), y(t, \omega)) : t \in T, x \in X, y \in Y, \omega \in \Omega\} \tag{41}$$

where X is the system input domain, Y is the output domain of the system, T represents a time domain, and Ω is an elementary events domain. There is a certain probabilistic, reason-result relationship between variables x and y , where x plays the role of a reason, and y - the result.

In paragraph 4.2 we considered the linguistic fuzzy model of the MISO system, assuming that x and y are linguistic variables (vector) with linguistic values determined by suitable fuzzy sets in the input and output numerical domain. Moreover, the probabilistic measure $p(x,y)$ on a set of realizations of the processes have been given. The model (29) can be treated as a joint probability of linguistic vector variable in the input-output domain.

Let us assume now, that the probabilistic measure $p(x,y)$ on a set of realizations of the processes $\{x(t), y(t)\}$, $t = t_k, k = 1, 2, \dots, K$ observed at the discrete moments, is given.

There are many models of stochastic systems discrete in a time domain T , for example an input-output dynamic model:

$$y(t_k) = f[x(t_k), x(t_{k-1}), \dots, x(t_{k-n}), \dots, y(t_{k-1}), \dots, y(t_{k-m})] \quad (42)$$

where $f()$ can be a multivariable regression function. These types of models are well known as Box-Jenkins' time series models and are modelled by using Takagi-Sugeno type fuzzy models (Yager, and Filev, 1994), (Hellendoorn, and Driankov, 1997).

We are interested in other types of models, which take into account a multivariable distribution function of the processes $\{x(t), y(t)\}$, $t = t_k, k = 1, 2, \dots, K$ observed at the discrete moments, e.g.

$$p(x, y) = p[x(t_k), x(t_{k-1}), \dots, x(t_{k-n}), \dots, y(t_k), y(t_{k-1}), \dots, y(t_{k-m})] \quad (43)$$

These models are used in more simple forms, as the first order models (e.g. white noise) or the second order models (e.g. Markov's process, known also as a *short memory process*).

The general form of the fuzzy model of a stochastic process discrete in a time domain T , can be expressed as a set of weighted rules:

$$\begin{aligned} &w_i [IF \ x(t_k) \text{ is } A_{i,k} \ \text{AND} \ x(t_{k-1}) \text{ is } A_{i,k-1} \ \dots \ \text{AND} \ x(t_{k-n}) \text{ is } A_{i,k-n} \ \dots \\ &\quad \text{AND} \ y(t_{k-1}) \text{ is } B_{i,k-1} \ \dots \ \text{AND} \ y(t_{k-m}) \text{ is } B_{i,k-m} \\ &\quad \text{THEN} \ y(t_k) \text{ is } B_{i,k}] \end{aligned} \quad (44)$$

where

$i=1, \dots, I$ - number of rules, determined by the partition of the input-output space $X^{n+1} \times Y^{m+1}$;

x, y -linguistic variables, $x \in X$, $y \in Y$ with linguistic values sets $L(X)$, $L(Y)$, determining linguistic states of the system,

$A_{i,k}, A_{i,k-1}, \dots, A_{i,k-n}$ - fuzzy subsets corresponding to linguistic values of variables $x(t_k), x(t_{k-1}), \dots, x(t_{k-n})$, $x \in X$;

$B_{i,k}, B_{i,k-1}, \dots, B_{i,k-m}$ - fuzzy subsets corresponding to linguistic values of variables $y(t_k), y(t_{k-1}), \dots, y(t_{k-m})$, $y \in Y$;

w_i - weight of i -th rule, a joint probability of the fuzzy event (fuzzy relation R_i) in the input-output space $X^{n+1} \times Y^{m+1}$ (according to (Walaszek-Babiszewska, 2007b))

$$w_i = P(R_i) = P(A_{i,k} \times A_{i,k-1} \times \dots \times A_{i,k-n} \times B_{i,k} \times B_{i,k-1} \times \dots \times B_{i,k-m}) \quad (45)$$

The weighted rule (44) can be easily written in a form of a rule with two weights, corresponding to a probability of the antecedent events and to a conditional probability of the consequent event, similarly to model (29).

4.4 Exemplary knowledge representation of a stochastic process

The data $\{x_{t_k}\}$ of the euro/Polish zloty exchange rate, observed daily in the first year of involving it into 12 countries of the EU, has been recognized as a realization of a certain stochastic process. The process has been modelled to predict some linguistic value of the process and the probability of its occurrence.

Two variables $x_{t_{k-1}}, x_{t_{k-2}}$ have been assumed as antecedent variables. From the point of view of fuzzy modelling, the created model represents a fuzzy relation $R(x_{t_k}, x_{t_{k-1}}, x_{t_{k-2}})$ of the linguistic variables in a form of weighted rules. Three linguistic states of the process have been distinguished: $L(X)=\{low(A_1), middle(A_2), high(A_3)\}$ and the fuzzy meaning have been defined, based on disjoint intervals in the process domain X .

The joint empirical probability distribution $p(x_{t_k}, x_{t_{k-1}}, x_{t_{k-2}})$ has been calculated, using disjoint cube intervals $a_i \times a_j \times a_k \in X^3, i,j,k=1,\dots,4$. The empirical probability distribution $P(x_{t_k}, x_{t_{k-1}}, x_{t_{k-2}})$ of linguistic variables, taking the fuzzy states A_1, A_2, A_3 observed at the moments t_k, t_{k-1}, t_{k-2} , has been computed. Then, the marginal and conditional probability distributions have been calculated.

The linguistic fuzzy model consists of 7 file rules. Table 3. presents all file rules of the model in a form of a decision table (Walaszek-Babiszewska, 2007b).

		x_{t-1}		
		A_1	A_2	A_3
x_{t-2}	A_1	$0.365 \begin{cases} x_t \text{ is } A_1 / 0.84 \\ x_t \text{ is } A_2 / 0.16 \\ x_t \text{ is } A_3 / 0 \end{cases}$	$0.06 \begin{cases} x_t \text{ is } A_1 / 0.42 \\ x_t \text{ is } A_2 / 0.55 \\ x_t \text{ is } A_3 / 0.03 \end{cases}$	0
	A_2	$0.07 \begin{cases} x_t \text{ is } A_1 / 0.4 \\ x_t \text{ is } A_2 / 0.43 \\ x_t \text{ is } A_3 / 0.17 \end{cases}$	$0.07 \begin{cases} x_t \text{ is } A_1 / 0.10 \\ x_t \text{ is } A_2 / 0.85 \\ x_t \text{ is } A_3 / 0.05 \end{cases}$	$0.025 \begin{cases} x_t \text{ is } A_1 / 0 \\ x_t \text{ is } A_2 / 0.48 \\ x_t \text{ is } A_3 / 0.52 \end{cases}$
	A_3	0	$0.03 \begin{cases} x_t \text{ is } A_1 / 0 \\ x_t \text{ is } A_2 / 0.30 \\ x_t \text{ is } A_3 / 0.70 \end{cases}$	$0.380 \begin{cases} x_t \text{ is } A_1 / 0 \\ x_t \text{ is } A_2 / 0.07 \\ x_t \text{ is } A_3 / 0.93 \end{cases}$

Table 3. The rule-based fuzzy model of the stochastic process $\{x_{t_k}\}$

5. References

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, Part I, Vol. 8, pp. 199-240

Zadeh, L.A. (1968). Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications*, Vol.23, No.2, pp. 421-427

Kacprzyk, J. (1986) *Fuzzy Sets in System Analysis* (in Polish), PWN, Warsaw

- Walaszek-Babiszewska, A. (2008). Probability measures of fuzzy events and linguistic fuzzy modelling -forms expressing randomness and imprecision, *Proc. WSEAS Conf. 'Art. Intell., Knowledge Eng., and Data Bases'*, pp. 207-212, Cambridge, UK, February 2008, WSEAS
- Walaszek-Babiszewska, A. (2004). *Statistical and Fuzzy Modeling of Grain Material Sampling and Operations; Selected Approaches*. Publishing House of the Silesian University of Technology, Gliwice, Poland
- Yager, R. R. & Filev, D.P. (1995). *Essentials of Fuzzy Modelling and Control* (in Polish), WNT and Wiley, Warsaw
- Walaszek-Babiszewska, A. (2007a). Construction of fuzzy models using probability measures of fuzzy events, *Proc.13th IEEE/IFAC Int.. Conf. on Methods and Models in Automation and Robotics, MMAR 2007*, pp. 661-667, August 2007, Szczecin, Poland
- Zadeh, L.A. (1979). A theory of approximate reasoning, *Machine Intelligence*, Vol.9, J.E Hayes, D. Michie and L.I. Mikulich, (Ed.), Elsevier, New York, pp. 149-194.
- Pedrycz, W. (1984). An identification algorithm in fuzzy relational systems. *Fuzzy Sets and Systems*, Vol.13, pp. 153-167
- Yager, R. R. & Filev, D.P. (1994). Template based fuzzy systems modelling. *Journal of Intelligent Systems*, Vol. 2, pp. 39-54
- Hellendoorn, H. & Driankov, D. (Ed.) (1997). *Fuzzy Model Identification; Selected Approaches*, Springer, Berlin - Heidelberg
- Walaszek-Babiszewska, A. (2007b). Linguistic knowledge representation for stochastic systems. *Proc. of the Int. Multiconference on Computer Science and Information Technology IMCSIT 2007 (in cooperation with IEEE), Wista, Poland; Vol. 2, pp. 141-150*



Automation and Robotics

Edited by Juan Manuel Ramos Arreguin

ISBN 978-3-902613-41-7

Hard cover, 388 pages

Publisher I-Tech Education and Publishing

Published online 01, May, 2008

Published in print edition May, 2008

In this book, a set of relevant, updated and selected papers in the field of automation and robotics are presented. These papers describe projects where topics of artificial intelligence, modeling and simulation process, target tracking algorithms, kinematic constraints of the closed loops, non-linear control, are used in advanced and recent research.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Anna Walaszek-Babiszewska (2008). Fuzzy Knowledge Representation Using Probability Measures of Fuzzy Events, Automation and Robotics, Juan Manuel Ramos Arreguin (Ed.), ISBN: 978-3-902613-41-7, InTech, Available from:

http://www.intechopen.com/books/automation_and_robotics/fuzzy_knowledge_representation___using_probability_measures_of_fuzzy_events

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.