Chapter 5
Stochastic-Fuzzy Knowledge-Based Approach to Temporal Data Modeling

Anna Walaszek-Babiszewska and Katarzyna Rudnik

Abstract. In the chapter an advanced fuzzy modeling method has been presented which can be useful in temporal data analysis. The method joints fuzzy and probabilistic approaches. The notions of the stochastic process with fuzzy states, and linguistic random variable have been defined to create a knowledge representation of the SISO and MISO dynamic systems. As the basic description of the stochastic process with fuzzy states observed at fixed moments, the joint probability distribution of n linguistic random variables has been assumed. The joint, conditional and marginal probability distributions of the stochastic process with fuzzy states valuate weights of particular rules of the knowledge rule base. Also, the probability distributions determine the probabilistic structure of the particular steps of the tested process. A mean fuzzy conclusion (prediction) can be calculated by the proposed inference procedure.

The implemented knowledge-based system, which creates the knowledge base with optimal number of elementary rules, has been also presented. The optimization method uses a fast algorithm to find fuzzy association rules as a process of automatic knowledge base extraction.

Two examples illustrate the presented methods of the knowledge base extraction from different numeric time series.

1 Introduction

In the topic literature there are many different approaches to time series modeling. The main distinguish can be made between statistical and fuzzy methods. Statistical methods are well known in econometrics and in control theory areas and there are many identification methods of the models. The key role in ordering of the statistical methods in time series modeling has played the work by Box and

Anna Walaszek-Babiszewska · Katarzyna Rudnik
Opole University of Technology, Opole, Poland
e-mail: {a.walaszek-babiszewska,k.rudnik}@po.opole.pl
a.walaszekbabiszewska@gmail.com

W. Pedrycz & S.-M. Chen (Eds.): Time Series Analysis, Model. & Applications, ISRL 47, pp. 97–118.
DOI: 10.1007/978-3-642-33439-9_5

Jenkins [2]. Also fuzzy approaches or more general, neuro-fuzzy and genetic-fuzzy approaches to time series modeling have a long history and a very large literature. Certain review of trends in fuzzy models and identification methods the interested reader can find e.g. in works [11] and [8].

It is often assumed, that temporal data collected from many objects of human activities constitute realizations of stochastic processes. Since the complete description of the stochastic process needs calculations of the series of *nD* probability distributions [3, 5, 12], many types of models have been invented, which are sufficient under the specific assumptions. Time series models are well known as the models of the specific realizations of time-discrete stochastic processes. In the fuzzy systems theory, the fuzzy representations of time-discrete stochastic processes are known in forms of the linguistic rule-based models, as well as, the Takagi-Sugeno-Kang (TSK) fuzzy models with equations at the consequent parts of rules [8, 10, 19].

In the chapter we present the method of temporal data analysis, which joints fuzzy and probabilistic approaches. The notions of the stochastic process with fuzzy states [18], and linguistic random variable have been defined. As the basic description of the stochastic process with fuzzy states observed at fixed moments $t_1, t_2, ..., t_n$, the joint probability distribution of n linguistic random variables has been assumed. The joint, conditional and marginal probability distributions of the stochastic process with fuzzy states determine respective weights of particular rules of the knowledge base. Also, the probability distributions are used to determine the probabilistic structure of the particular steps of the tested process and to calculate a mean fuzzy conclusion (prediction) by the proposed inference procedure.

We also present the implementation of the knowledge-based system [13], which creates the probabilistic-fuzzy knowledge base with the optimal number of elementary rules. The optimization method uses a fast algorithm to find fuzzy association rules as a process of automatic knowledge base extraction.

Exemplary calculations are presented with results derived by using the implemented knowledge-based system and chosen numeric time series.

2 Stochastic Process with Fuzzy States

2.1 Introduction

According to the theory of stochastic processes, a family of time dependent random variables (dependent on a real parameter *t*), denoted as

$$\{X(t,\omega),\ X \in \chi, t \in T, \omega \in \Omega\}, \tag{1}$$

is defined as stochastic process (shortly written as $X(t)$), where $\chi \subseteq R$ is a domain of the process values, $T \subset R$ is a domain of parameter *t*, and Ω is an elementary events domain.

For each given t, $X(t) = X_t$ is a random variable and $F_t(x) = P\{X_t < x\}$ is a distribution function of X_t.

For any given set of parameter values, $\{t_1, t_2, ..., t_n\}$, stochastic process $X(t)$ is determined by *n-D* probability distribution function

$$F_{t_1,t_2,...,t_n}(x_1, x_2,..., x_n) = P\{X(t_1) < x_1,..., X(t_n) < x_n\}$$
$$= P\{X_{t_1} < x_1,..., X_{t_n} < x_n\}. \quad (2)$$

Stochastic process is fully determined by a family of all *n-D* probability distribution functions, where n=1,2,... For any given elementary event $\omega' \in \Omega$, the function $x(t) = X(t, \omega')$ is a realization (trajectory) of the stochastic process $X(t)$ [3, 5, 12].

2.2 One Dimensional Probability Distribution of the Stochastic Process with Fuzzy States

Let $X(t)$ denotes a stochastic process, a family of time dependent random variables, taking its values in $\chi \subset R$, $t \in T \subset R$. Let (χ, $\mathcal{B}$, p) be a probability space, where $\mathcal{B}$ is a σ-field of Borel sets in $\chi \subset R$ and p is a probability measure over (χ, $\mathcal{B}$).

Let us determine, in the domain of the stochastic process values χ, a linguistic variable which is generated by the process $X(t)$, at fixed t. The linguistic variable is given by quintuple $< X_t, L(X), \chi, G, M >$, where X_t is the name of the variable and $L(X)=\{LX_i\}$, i=1,2,...,I is a collection of its linguistic values. The semantic rule M assigns fuzzy event A_i, i=1,2,...,I to every meaning of LX_i, i=1,2,...,I [21]. Let also, membership functions $\mu_{A_i}(x): \chi \to [0,1]$ be Borel measurable and meet

$$\sum_{i=1}^{I} \mu_{A_i}(x) = 1, \ \forall x \in \chi. \quad (3)$$

Then, the collection of linguistic values $L(X)=\{LX_i\}$, i=1,2,...,I and the collection of corresponding fuzzy sets A_i, $i=1,2,...,I$ defined over χ, will be called the *linguistic (fuzzy) states of the stochastic process X*(t).

According to Zadeh's definitions from [20], fuzzy states A_i, $i=1,2,...,I$ of the stochastic process constitute *fuzzy events* in the probability space (χ, $\mathcal{B}$, p). Probability of the occurrence the fuzzy state A_i, can be calculated by the following Lebesgue-Stietljes' integral

$$P(A_i) = \int_{x \subseteq \chi} \mu_{A_i}(x) dp, \quad (4)$$

if the integral exists [20]. The existence of the integral (4) results from the assumption that $\mu_{Ai}(x)$ is a Borel measurable function. If the universal set is a countable collection, $\mathcal{X}=\{x_n\}$, $n=1,2,\ldots$, and the probability function is determined for discrete process values $P(X=x_n)=p_n$, such that $\sum_n p_n = 1$, then the probability of fuzzy event $A_i = \sum_n (\mu_{Ai}(x_n)/x_n)$, denoted as $P(A_i)$, is defined as

$$P(A_i) = \sum_n \mu_{A_i}(x_n) p_n \,. \tag{5}$$

One dimensional probability distribution of linguistic values (fuzzy states) of the stochastic process $X(t)$, for any fixed value t, can be defined as a set of probabilities of fuzzy events

$$P(X_t) = \{P(A_i\}, i = 1,2,\ldots, I \,, \tag{6}$$

where $P(A_i)$, i=1,2,...,I are determined according to (4) or (5) and the following relationships must be fulfilled [16]:

$$0 \le P(A_i) \le 1,\ i=1,2,\ldots,I \ ; \qquad \sum_{i=1}^{I} P(A_i) = 1 \ . \tag{7}$$

2.3 nD Probability Distribution of the Stochastic Process with Fuzzy States

One-dimensional probability of the stochastic process is an efficient description for the special type of stochastic processes, so called 'white noise processes'.

To determine a probability description of the stochastic process with fuzzy states for two fixed moments t_1, t_2, let us take into account two random variables $(X(t_1), X(t_2))$, determined in the probability space ($\mathcal{X}^2$, $\mathcal{B}$, p), where $\chi^2 \subseteq R^2$. Two linguistic random variables (linguistic random vector) (X_{t_1}, X_{t_2}) generated by stochastic process values in χ^2, can be defined. The simultaneous linguistic values $LX_i \times LX_j$, *i,j=1,2,…,I* and the corresponding collection of fuzzy events $\{A_i \times A_j\}_{i,j=1,\ldots,I}$ can be determined over χ^2 by membership functions $\mu_{A_i \times A_j}(u): \chi^2 \to [0,1]$, *i,j=1,2,…,I.* The membership functions $\mu_{A_i \times A_j}(u)$ in the linguistic vector domain, χ^2, should be Borel measurable and fulfill the following relationship:

$$\sum_{i=1}^{I}\sum_{j=1}^{I} \mu_{A_i \times A_j}(u) = 1, \ \forall (u) \in \chi^2 . \tag{8}$$

Then, fuzzy sets $A_i \times A_j$ determined in χ^2 are the *simultaneous fuzzy events* and probability $P(A_i \times A_j)$, is defined according to (4), as follows

$$P(A_i \times A_j) = \int_{u \in \chi^2} \mu_{A_i \times A_j}(u) dp , \tag{9}$$

where

$$\mu_{A_i \times A_j}(u) = T(\mu_{A_i}(x), \mu_{A_j}(x)) , \tag{10}$$

in particular

$$\mu_{A_i \times A_j}(u) = \mu_{A_i}(x)\mu_{A_j}(x) . \tag{11}$$

If universe $\mathcal{X}^2$ is a finite set, $\chi^2 = \{(x_k, x_l)\}$, k=1,…,K, l=1,…,L, then the *probability of simultaneous fuzzy event* $A_i \times A_j$ is determined, according to

$$P(A_i \times A_j) = \sum_{(x_k, x_l) \in \chi^2} p(x_k, x_l) \mu_{A_i \times A_j}(x_k, x_l) , \tag{12}$$

where $\{p(x_k, x_l)\}_{k=1,2,...k; l=1,2,...,L}$ is a probability function of the discrete random vector variable $(X(t_1), X(t_2))$, at two fixed moments t_1, t_2 .

The *joint 2D probability distribution of the linguistic values (fuzzy states) of the stochastic process* $X(t)$ is determined by the collection of probabilities of *simultaneous fuzzy events* $A_i \times A_j$

$$P(X_{t_1}, X_{t_2}) = \{P(A_i \times A_j)\}_{, i, j=1,2,...,I} , \tag{13}$$

if the following relationships are fulfilled

$$0 \le P(A_i \times A_j) \le 1, \ \forall i, j = 1,..., I \ \text{ and } \ \sum_{i=1}^{I}\sum_{j=1}^{I} P(A_i \times A_j) = 1 . \tag{14}$$

To determine the *nD* probability distribution of the stochastic process with fuzzy states, assume first, that stochastic process $X(t)$, for a set of moments $t_1,...,t_n$ is represented by a random vector $(X(t_1),..., X(t_n))$ and $(\mathcal{X}^n, \mathcal{B}, p)$ is a probability space. Let the linguistic variables

$$< X_{t_1}, L(X), \chi, G, M >,...,< X_{t_n}, L(X), \chi, G, M > \tag{15}$$

be generated by the stochastic process in the domain χ. The same sets of the linguistic values

$$L(X_{t_1}) = ... = L(X_{t_n}) = L(X) = \{LX_i\}_{i=1,...,I}\,, \tag{16}$$

for particular linguistic variables are represented by fuzzy sets A_i, i=1,…,I, with membership functions, $\mu_{A_i}(x) : \chi \to [0,1]$.

Let the random linguistic vector variable whose name is determined by a vector $\left(X_{t_1},...,X_{t_n}\right)$ *takes simultaneous linguistic values*

$$L(X)^n = \{LX_{i_1} \times LX_{i_2} \times ... \times LX_{i_n}\},\ \forall i_1,...,i_n = 1,...,I\,, \tag{17}$$

whose meanings are represented by the collection of *simultaneous fuzzy events* (*fuzzy states*)

$$\{\ (A_{i_1} \times ... \times A_{i_n})\ \},\ \forall i_1,...,i_n = 1,...,I\,. \tag{18}$$

Fuzzy events (18) are determined on χ^n by membership functions $\mu_{A_1{}_i \times ..\times A_{i_n}}(u),\ u \in \chi^n$, which are Borel measurable and fulfill the relationship

$$\sum_{i_1=1}^{I} ... \sum_{i_n=1}^{I} \mu_{A_{i1} \times ..\times A_{i_n}}(u) = 1, \forall u \in \chi^n. \tag{19}$$

Let also probabilities of the simultaneous fuzzy events (18), calculated according to (4) or (5), respectively, exist and fulfill the relationships

$$0 \le P(A_{i_1} \times ... \times A_{i_n}) \le 1,\ \ \forall i_1,...,i_n = 1,...,I\,; \tag{20}$$

$$\sum_{i_1=1}^{I} ... \sum_{i_n=1}^{I} P(A_{i_1} \times ... \times A_{i_n}) = 1\,. \tag{21}$$

Then, *nD joint probability distribution of linguistic values (fuzzy states) of the stochastic process X(t) at moments* $t_1,...,t_n$ is a probability distribution of linguistic vector variable $\left(X_{t_1},...,X_{t_n}\right)$, determined by the following collection of probabilities of the simultaneous fuzzy events [16]

$$P(X_{t_1},...,X_{t_n}) = \{P(A_{i_1} \times ... \times A_{i_n})\}_{i_1=1,...,I;...;i_n=1,...,I}\,. \tag{22}$$

In the n*D* joint probability distribution of linguistic values of the stochastic process $X(t)$ we can distinguish *rD* , $r<n$ marginal probability distributions, e.g.

$$P(X_{t_1},...,X_{t_{n-1}}) = \{\sum_{i_n=1}^{I} P(A_{i_1} \times ... \times A_{i_{n-1}} \times A_{i_n})\}_{i_1=1,...,I;...;i_{n-1}=1,...,I}\,, \tag{23}$$

as well as, conditional probability distributions. Generally, *nD* joint probability distribution of linguistic values of the stochastic process $X(t)$ can be expressed by using marginal and conditional probability distributions in a way

$$P(X_{t_n}, X_{t_{n-1}}, ..., X_{t_2}, X_{t_1}) = P(X_{t_n} / X_{t_{n-1}}, ..., X_{t_2}, X_{t_1}) P(X_{t_{n-1}}, ..., X_{t_2}, X_{t_1}) = \\ P(X_{t_n} / X_{t_{n-1}}, ..., X_{t_2}, X_{t_1}) P(X_{t_n-1} / X_{t_{n-2}}, ..., X_{t_2}, X_{t_1}) .. P(X_{t_2} / X_{t_1}) P(X_{t_1}). \tag{24}$$

2.4 Fuzzy Mean Value of the Stochastic Process with Fuzzy States

Let the stochastic process $X(t)$ takes its linguistic values $L(X)=\{LX_i\}$, $i=1,2,\ldots,I$, which are represented by fuzzy events A_i, , $i=1,2,\ldots,I$ in χ. Let the probability distribution of the fuzzy states, $P(X_t) = \{P(A_i\}, i = 1,2,..., I$ exists. Then, a *fuzzy mean value of the stochastic process with fuzzy states*, denoted as $\overline{A}(X)$, is a fuzzy set determined as

$$\overline{A}(X) = \sum_{i=1}^{I} A_i P(A_i), \ \forall x \in \chi, \tag{25}$$

and the membership function is calculated as follows:

$$\mu_{\overline{A}}(x) = \sum_{i=1}^{I} \mu_{A_i}(x) P(A_i), \ \forall x \in \chi. \tag{26}$$

3 Fuzzy Knowledge Base of the Stochastic Systems

Fuzzy rule based models of dynamic systems are being used not only when knowledge about the real system functioning is incomplete but also when the fuzzy rule based model has to approximate the real system characteristics when the system is too complex or nonlinear. Those models are well known and they are described in the subject literature, eg. in [8, 19]. They are often connected with algorithms of clustering or evolving algorithms.

The novelty in the propositions implemented into model known in subject literature through this work, is the model validation by the probability distributions determined by empirical data.

Defining the fuzzy knowledge base for stochastic environment, it is necessary to make some assumptions about the possibility of existing multidimensional probability distributions of stochastic processes realizations observed in long time intervals. Usually, we assume also ergodicity and stationarity of the processes.

3.1 Fuzzy SISO Model of the Stochastic Process

Let $X(t)$ be a stochastic process with fuzzy states, as it was shown in paragraph 2. Assuming that the process was observed at two fixed moments $t_1, t_2 \in T;\ t_2 > t_1$, the process realizations have been used to calculate 2*D* empirical probability distributions of fuzzy states.

The fuzzy knowledge representation of the stochastic process is a collection of the following weighted file rules, in the form [18]:

$$\forall A_i \in L(X),\ i=1,\ldots,I$$

$$R^{(i)}: w_i[\mathit{If}\ (X_{t_1}\ \mathit{is} A_i)]\ \mathit{Then}\, (X_{t_2}\ \mathit{is}\ A_1)w_{1/i}$$

$$-----------$$

$$\mathit{Also}\,(X_{t_2}\ \mathit{is}\ A_j)w_{j/i} \quad , \tag{27}$$

$$-----------$$

$$\mathit{Also}\,(X_{t_2}\ \mathit{is}\ A_J)w_{J/i}$$

where $A_i, A_j \in L(X), i.j = 1,2,...,I$ denote the fuzzy states of the process, and weights

$$w_i = P(X_{t_1} = A_i),\ i=1,2,\ldots,I \tag{28}$$

are the probabilities of fuzzy events at the antecedents of the rules (marginal probability distribution), and weights

$$w_{j/i} = P[(X_{t_2} = A_j)/(X_{t_1} = A_i)],\ j=1,2,\ldots,I;\ i=const \tag{29}$$

are the conditional probabilities of the fuzzy events at the consequent part of the rules (conditional probability distribution). According to the probability distribution features, the following relationships are fulfilled

$$\sum_{i=1,\ldots,I} w_i = 1 \quad \sum_{j=1,\ldots,J} w_{j/i} = 1\,.$$

The model can be also presented as a collection of the elementary weighted rules

$$\forall A_i \in L(X),\ \forall A_j \in L(X),\ i,j=1,2,\ldots,I$$

$$R^{(i,j)}: w_{ij}[\mathit{If}\ (X_{t_1}\ \mathit{is}\ A_i)\ \mathit{Then}\ (X_{t_2}\ \mathit{is}\ A_j)], \tag{30}$$

where

$$w_{ij} = P(X_{t_1}, X_{t_2}) = P(A_i \times A_j),\ i,j=1,2,\ldots,I \tag{31}$$

is a joint probability of fuzzy events in the rule (joint probability distribution) and the following relationship must be fulfilled

$$\sum_{i=1,\dots,I}\ \sum_{j=1,\dots,J} w_{ij} = 1 .$$

The above propositions of the knowledge representation contain weights: w_i, $w_{j/i}$, w_{ij}, which stand for the frequency of the occurrence the fuzzy events in particular parts of rules. The weights, real numbers from the interval [0, 1], do not change logic values of the sentences.

Assuming stationarity of the process with fuzzy states, the prediction of the process can be determined by means of approximate reasoning.

3.2 Inference Procedure (Prediction Procedure) from the SISO Model of the Stochastic Process

For the logic analysis we take into account the following fuzzy relation representing file rule (30)

$$R^{(i)} : A_i \Rightarrow (A_1 \cup \dots \cup A_j \cup \dots \cup A_I)\,, \tag{32}$$

which can be described by membership function

$$\mu_{R^{(i)}}(x_{t_1}, x_{t_2}) = \mu_{A_i \Rightarrow (A_1 \cup \dots \cup A_I)}(x_{t_1}, x_{t_2})\,,\ i=1,\dots,I. \tag{33}$$

To consider the *prediction procedure*, which is based on well-known procedure of approximate reasoning (e.g. in [8, 10, 19]), let us assume the crisp value of the stochastic process at moment t_1, $X(t_1) = x^*_{t_1}$.

Then the level of activation of the elementary rule is determined as

$$\tau_i = \mu_{A_i}(x^*_{t_1})\,,\ i=1,\dots,I \tag{34}$$

and the fuzzy value of the conclusion $(X_{t_2}\ is\ A'_{j/i})$, computed e.g. based on Larsen's rule of reasoning, is a fuzzy set $A'_{j/i}$, determined by its membership function

$$\mu_{A'_{j/i}}(x_{t_2}) = \tau_i \mu_{A_j}(x_{t_2})\,,\ j=1,\dots,I;\ i=\text{const}. \tag{35}$$

The fuzzy conditional expected value (fuzzy conditional mean value) of the output of *i-th* rule $E\{(X_{t_2}\ is\ \varphi(A_j))/[X(t_1)\ is\ A_i]\} = A'_i$, stands for the aggregated outputs of elementary rules, $j=1,\dots,J$, according to the formula [16]

$$\mu_{A'_i}(x_{t_2}) = \sum_j w_{j/i}\,\mu_{A'_{j/i}}(x_{t_2})\,. \tag{36}$$

The fuzzy expected value of the prediction, $E\{(X_{t_2}\ is\,\varphi(A_j))/[X(t_1)=x_{t_1}*]\}=A'$, computed as the aggregated outputs of all active *i-th* rules, is determined by the formula [16]

$$\mu_{A'}(x_{t_2})=\sum_i w_i\mu_{A_i^{'}}(x_{t_2})=\sum_i w_i\tau_i\sum_j w_{j/i}\mu_{A_j}(x_{t_2})\,. \tag{37}$$

The prediction according to the generalized Mamdani-Assilian's type interpretation of fuzzy models gives us the following conclusion

$$\mu_{A_{j/i}^{'}}(x_{t_2})=T(\tau_i,\mu_{A_j}(x_{t_2}))\,,\ j=1,\ldots,I;\ i=\text{const}. \tag{38}$$

Prediction determined by using the logic type interpretation of fuzzy models, gives us the following relationships, instead of (35) or (38):

$$\mu_{A_{j/i}^{'}}(x_{t_2})=I(\tau_i,\mu_{A_j}(x_{t_2}))\,,\ j=1,\ldots,I;\ i=const, \tag{39}$$

where T denotes a t-norm and I means the implication operator.

The scheme of the prediction procedure from the SISO model is presented in Fig. 1.

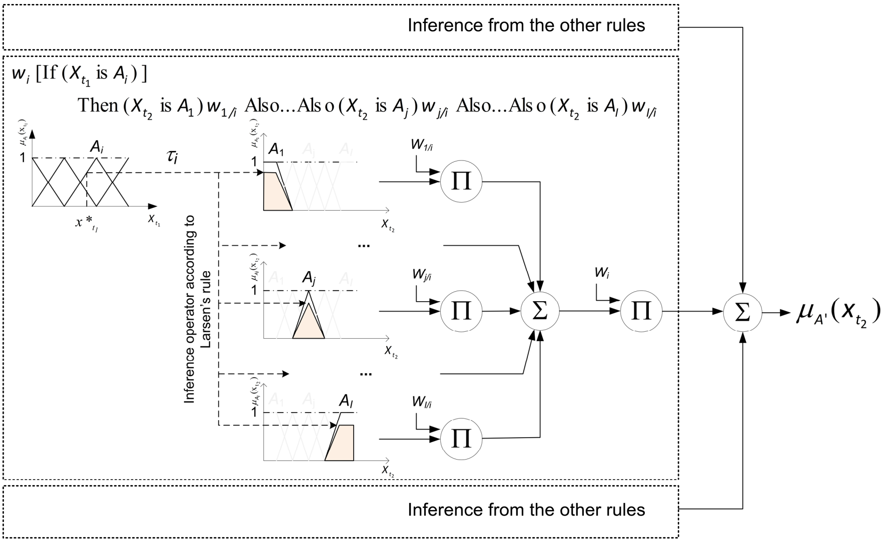

Fig. 1 Scheme of the prediction procedure from the SISO fuzzy model of the stochastic process

3.3 Fuzzy MISO Model of the Long Memory Stochastic Process

Let the stochastic process *X(t)* with fuzzy states be determined, as it has been shown in paragraph 2. To create the representation of knowledge base in the form of fuzzy *If…Then* rules, the linguistic random variables have been determined with their collection of linguistic values, $\{LX_{k,i_k}\}$, the same for each variable. The meanings of linguistic values are represented by fuzzy sets, A_{k,i_k}, $k=1,\dots,n;$ $i_k=1,\dots,I$ in χ, with the Borel measurable membership functions $\mu_{A_{k,i_k}}(x):\chi\to[0,1]$, $k=1,\dots,n;$ $i_k=1,\dots,I$.

The fuzzy sets divide the space χ^n of values of samples into *n-D* fuzzy areas: $(A_{1,i_1}\times\dots\times A_{n,i_n})$, $i_1=1,2,\dots,I;\dots;\ i_n=1,2,\dots,I$.

Assuming that the process was observed at fixed moments $t_1,t_2,\dots,t_n\in T;\ t_n>t_{n-1}>\dots>t_1$, the process realizations have been used to calculate the following empirical probability distributions of fuzzy states:

- *nD* joined probability distribution of the linguistic random vector variable $\left(X_{t_1},\dots,X_{t_n}\right)$

$$P(X_{t_1},\dots,X_{t_n})=\{P(A_{1,i_1}\times\dots\times A_{n,i_n})\}_{i_1=1,\dots,I;\dots;i_n=1,\dots,I}\,, \tag{40}$$

- marginal $(n\text{-}1)D$ probability distribution of the linguistic random vector variable (of the antecedent fuzzy events)

$$w_j=P(X_{t_1},\dots,X_{t_{n-1}})=\{P(A_{1,i_1}\times\dots\times A_{n-1,i_{n-1}})\}_{i_1=1,\dots,I;\dots;i_{n-1}=1,\dots,I}\,, \tag{41}$$

- conditional probability distribution (of the consequent fuzzy events)

$$w_{i_n/j}=P(X_{t_n}/X_{t_1},\dots,X_{t_{n-1}})=\{P[A_{n,i_n}/(A_{1,i_1}\times\dots\times A_{n-1,i_{n-1}})]\} \tag{42}$$

$$i_n=1,2,\dots,I;\ i_1,i_2,\dots,i_{n-1}=const.$$

The MISO fuzzy model, as the knowledge representation of the stochastic process, has the form of the collection $\{R^{(j)}\}_{j=1,2,\dots,J}$ of weighted file rules [16]:

$$\forall A_{1,i_1}\in L(X_{t_1}),\ \forall A_{2,i_2}\in L(X_{t_2}),\dots,\forall A_{n,i_n}\in L(X_{t_n}),$$
$$i_1,i_2,\dots,i_n=1,2,\dots,I;$$

$$\begin{aligned}
R^{(j)}: w_j[&If\ (X_{t_1}\ is\ A_{1,i_1})\,And(X_{t_2}\ is\ A_{2,i_2})\,And\dots And(X_{t_{n-1}}\ is\ A_{n-1,i_{n-1}})]\\
&Then(X_{t_n}\ is\ A_{n,1})w_{1/j}\\
&------------\\
&Also(X_{t_n}\ is\ A_{n,i_n})w_{i_n/j}\\
&------------\\
&Also(X_{t_n}\ is\ A_{n,I})w_{I/j}
\end{aligned} \tag{43}$$

j=1,2,…J ; J - number of file rules. The number of file rules, J, depends on the number of fuzzy (n-1)D areas in the input space $\mathcal{X}^{n-1}$, with non-zero probabilities w_i ; it can be even $J = I^{n-1}$.

3.4 Prediction Procedure from the MISO Fuzzy Model of the Stochastic Process

Assuming, that the process with fuzzy states is a stationary process, that is, the *nD* joint probability distribution does not depend on time, we can use the created knowledge representation for the prediction of the process. The conclusion, fuzzy or numeric, determined by means of approximate reasoning represents the prediction of the process. The input data can be fuzzy or numeric in their character.

Let us consider the prediction procedure, assuming crisp data of observations of the process, $X(t_1) = x^*_{t_1}, X(t_2) = x^*_{t_2}, ..., X(t_{n-1}) = x^*_{t_{n-1}}$. Then, the level of activation of *j-th* rule is determined by the t-norm of membership functions of fuzzy sets in antecedents as follows [8, 10, 19]:

$$\tau_j = T_1\left(\mu_{A_{1,i_1}}(x^*_{t_1}), \mu_{A_{2,i_2}}(x^*_{t_2}), ..., \mu_{A_{n-1,i_{n-1}}}(x^*_{t_{n-1}})\right), j=1,\dots,J. \tag{44}$$

If the values of the process at moments $t_1, t_2, ..., t_{n-1}$ are expressed by fuzzy numbers (linguistic values), that is

$$(X_{t_1} \, is \, A'_{1,i_1}) \; And \; (X_{t_2} \, is \, A'_{2,i_2}) \, And \, \dots And \; (X_{t_{n-1}} \, is \, A'_{n-1,i_{n-1}}),$$

where A'_{k,i_k} are given by the membership functions $\mu_{A'_{k,i_k}}(x_{t_k}) : \chi \to [0,1]$, k=1,2,…,n-1, then the level of activation of *j-th* rule is expressed as [8, 10, 19]:

$$\tau_j = T_1\left\{\left\{\sup_{x\in\chi}[\mu_{A'_{1,i_1}}(x_{t_1}) \wedge \mu_{A_{1,i_1}}(x_{t_1})]\right\}, ..., \left\{\sup_{x\in\chi}[\mu_{A'_{n-1,i_{n-1}}}(x_{t_{n-1}}) \wedge \mu_{A_{n-1,i_{n-1}}}(x_{t_{n-1}})]\right\}\right\}. \tag{45}$$

The *fuzzy conclusion*, $X_{t_n} \, is \, A'_{n,i_n}$, from i_n-*th* consequent part of the rule can be determined by one of the ways [8, 10, 19]:

- according to Mamdani-Assilian's rule of inference

$$\mu_{A'_{n,i_n}}(x_{t_n}) = \tau_j \wedge \mu_{A_{n,i_n}}(x_{t_n}), \tag{46}$$

- according to Larsen's rule

$$\mu_{A'_{n,i_n}}(x_{t_n}) = \tau_j \mu_{A_{n,i_n}}(x_{t_n}), \tag{47}$$

- according to generalized Mamdani-Assilian's type of interpretation

$$\mu_{A'_{n,i_n}}(x) = T_2\left(\tau_j, \mu_{A_{n,i_n}}(x)\right) \tag{48}$$

- according to the logic interpretation

$$\mu_{A'_{n,i_n}}(x_{t_n}) = I\left(\tau_j, \mu_{A_{n,i_n}}(x_{t_n})\right). \tag{49}$$

The fuzzy conditional expected value of the conclusion A'_{n,i_n}

$$E\left[\left(X_{t_n} is\, A'_{n,i_n}\right)/\left(X_{t_1} is\, A'_{1,i_1}\right)\cap\ldots\cap\left(X_{t_1} is\, A'_{1,i_1}\right)\right] = A'_{n/j} \tag{50}$$

is the aggregated value (weighted sum) of conclusions from particular i_n*-th* outputs, i_n=1,…,I (calculated according to one of relationships (46) - (49)) and the conditional probabilities of fuzzy events in i_n*-th* consequents, as follows

$$\mu_{A'_{n/j}}(x_{t_n}) = \sum_{i_n} w_{i_n/j}\mu_{A'_{n,i_n}}(x_{t_n}). \tag{51}$$

Taking into account all *j-th* active rules, the *fuzzy conditional expected value of the prediction*, A_n',

$$A_n' = E\left\{(X_{t_n} is\, A'_{n/j})/\left(X(t_1) = x^*_{t_1}, X(t_2) = x^*_{t_2}, \ldots, X(t_{n-1}) = x^*_{t_{n-1}}\right)\right\} \tag{52}$$

can be calculated as the aggregated value (weighted sum) of conclusions from particular *j-th* file rules, $A'_{n/j}$, (51), and joint probabilities of fuzzy events in particular antecedents, as follows

$$\mu_{A'_n}(x_{t_n}) = \sum_j w_j \mu_{A'_{n/j}}(x_{t_n}) = \sum_j w_j \sum_{i_n} w_{i_n/j}\mu_{A'_{n,i_n}}(x_{t_n}). \tag{53}$$

Subscript n in A_n' shows, that fuzzy conclusion is the fuzzy value of the linguistic random variable X_{t_n}.

The discussed prediction procedure from the MISO model is presented in Fig. 2.

The numerical value of the prediction, x^*_n, can be determined as the centroid of A_n' calculated e.g. by the COA method

$$x^*_n = \frac{\int_{x\in\chi} x\mu_{A'_n}(x)dx}{\int_{x\in\chi} \mu_{A'_n}(x)dx}. \tag{54}$$

The fuzzy value of the prediction $(X_{t_n} is A_n^{'})$ is a function of fuzzy propositions $(X_{t_k} is A'_{k,i_k})$, or numerical propositions $X(t_k) = x_{t_k}^*$, for k=1,2,…,n-1, on the input of the system, as well as the chosen procedures of inference.

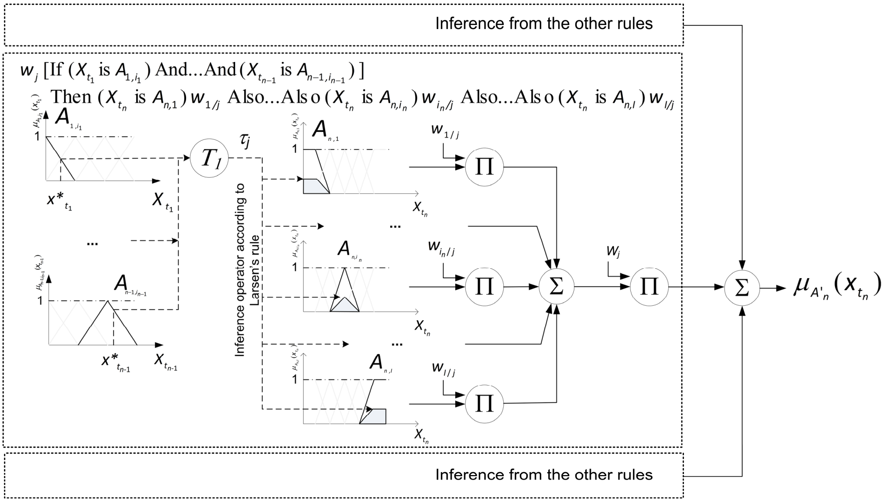

Fig. 2 Scheme of the prediction procedure from the MISO fuzzy model of the stochastic process

3.5 Probability of Fuzzy Predictions

We can also determine the probability of the fuzzy conclusions, fuzzy predictions, derived from the stochastic-fuzzy rule bases of the SISO and the MISO model.

Since the fuzzy conclusion (prediction), determined during the reasoning procedures, is given by its membership function in a domain of the output variable, and the probability distribution $p(x)$ has been determined based on data, then, probability of the fuzzy prediction can be determined by the following formula

$$P\Big((X_{t_n} \; is \, A_n^{'})/(X(t_1)=x_{t_1}^*,...,X(t_{n-1})=x_{t_{n-1}}^*)\Big) = \int_{x\in\chi} \mu_{A_n^{'}}(x)p(x)dx \, . \quad (55)$$

4 Conception of the Knowledge-Based Inference System

The knowledge-based systems are usually composed of the following parts [8, 10, 19]:

- knowledge base in the form of if-then rules (43), that contains information essential to solve a given problem,

- fuzzification block that transforms quantitative data into qualitative data represented by fuzzy sets on the bases of membership grades entered in the database,
- inference block that utilizes the database and the implemented aggregation methods and final inference (reasoning) to solve specialized problems,
- defuzzification block that calculates the crisp value (defuzzified value) at the system output on the bases of the resulting membership grades.

The implemented knowledge base contains database and rule base. The database contains information defined by experts on a given application field containing linguistic values of the variables accounted in the rule base and definitions of fuzzy sets identified with these values. On the other hand, knowledge base contains a set of linguistic rules created on the grounds of a modified algorithm generating fuzzy association rules. The algorithm makes it possible to adjust the model to measurement data. The characteristic form of the rules, exposing an empirical probability distribution of fuzzy events enables a simple interpretation of the knowledge contained in the model and additional analysis of the considered problem.

The inference mechanism with multiple inputs and a single output enables the calculation of the membership function of the conclusion, on the bases of the crisp input data, and, in consequence, the defuzzified value of the model output. For the system with the rule base in the form of (43), there are many possible ways of obtaining crisp output results. In this conception of the knowledge-based inference system, we consider the methods presented in chapter 3.4.

4.1 Methods of Fuzzy Knowledge Discovery

The *if-then* rules that constitute the knowledge bases of the fuzzy system may be defined in two ways:

- as logical rules constituting subjective definitions created by experts on the grounds of experience and knowledge of the investigated phenomenon,
- as physical rules constituting objective knowledge models defined on the grounds of observations and natural research into the analyzed process (object) and its regularities.

In the case of fuzzy modeling there were initially logics rules, yet, in consideration of machine learning a hybrid of rules was gradually implemented according to which initial assumptions concerning fuzzy sets and the associated rules are defined following the experts' conviction, whereas other parameters are adjusted to measurement data. The objective of automatic data discovery is to obtain the smallest set of *if-then* rules enabling as accurate representation of the modelled object or phenomenon as possible.

Methods of knowledge discovery for fuzzy systems of Mamdani type include [10, 19]:

- Wang-Mendel method,
- Nozaki-Ishibuchi-Tanaki method,
- Sugeno-Yasukawa method,
- template-based method of modelling fuzzy systems.

In order to obtain databases for fuzzy systems, data mining methods have also been applied.

Data mining, considered as the main stage in knowledge discovery [4] is focused on non-trivial algorithms of searching "hidden", so far unknown and potentially required information [6] and its records in the form of mathematical expressions and models. Some of the data mining methods identify zones in the space of system variables, which, consequently, create fuzzy events in the rules. This may be accomplished by searching algorithm clusters or covering algorithms, also called separate and conquer algorithms. Other methods, for example: fuzzy association rules, are based on constant division for each attribute (fuzzy grid) and each grid element is regarded as a potential component of the rule. As far as the first approach is concerned, each identified rule has its own fuzzy sets [17]. Therefore, from the point of view of rules interpretation, the second approach seems more applicable [9].

4.2 Association Rules as Ways of Fuzzy Knowledge Discovery

Irrespective of automatic knowledge discovery, rules of the fuzzy model are obtained on the bases of their optimal adjustment to experimental data. In view of this, the generation of the rules may be understood as a search for rules with high occurrence frequency, where, the frequency parameter influences the optimal rules adjustment. In such case, fuzzy rules may be analyzed as the co-existence of fuzzy variable values in experimental data, i.e.: fuzzy association rules.

The issue of association rules was first discussed in [1]. Nowadays it is one of the most common data mining methods. In a formal approach, the association rules have the form of the following implications:

$$X \Rightarrow Y\ (\mathrm{s,c}), \tag{56}$$

where X and Y are separable variable sets (attributes) in the classic approach to mathematical sets, often referred to as: X – conditioning values set, Y- conditioned values set.

Considering the fuzzy rules of association for the MISO model (43), the following may be derived:

$$X \Rightarrow Y:$$
$$(X_{t_1}\ is\ A_{1,i_1})And(X_{t_2}\ is\ A_{2,i_2})And...And(X_{t_{n-1}}\ is\ A_{n-1,i_{n-1}}) \Rightarrow X_{t_n}\ is\ A_{n,1}\ (s,c), \tag{57}$$

where A_{k,i_k} , $k=1,\ldots,n;\ i_k=1,\ldots,I$ denote the fuzzy states of the process.

Each association rule is connected with two statistical measures that determine the validity and power of the rule: *support* (s) – probability of the simultaneous incidence of set ($X \cap Y$) in the set collection and *confidence* (c) – also called credibility which is conditional probability ($P(Y \mid X)$). The issue of discovering fuzzy association rules involves finding, in a given database, all support and trust values that are higher than the association rules the support and trust of which are higher than the defined minimal values of support and trust given by users.

The first application of the association rules was in basket analysis. However, taking into account the fact that rules may include variables that are derived from diverse variables expressed in a natural language, the ranges of the application of the discussed method may be extended to forecasting, decision-making, planning, control etc. In the inference system with stochastic-fuzzy knowledge base, we proposed to use the idea of fuzzy association rules to knowledge discovery.

In the topic literature we can find many algorithms of creating the association rules and modifications [1, 7] but they generate association rules only in non-fuzzy version. To knowledge discovery in the form of (43) two algorithms have been proposed [13, 14, 17]. One is based on the *Apriori* algorithm and the second algorithm uses the *FP-Growth* assumption. In these algorithms the so called frequent fuzzy set is a set of which the probability of the occurrence is bigger than the value of the assumed minimal support s. Thus, the inputs of the proposed algorithm are: set of measurements used for model identification, predefined database (linguistic values of variables considered in the model and definitions of fuzzy sets identified with the values), and the threshold value of minimal support (s). Threshold value of the minimal confidence (c) is not in use. The output of the algorithms is a rule base of a probabilistic-fuzzy knowledge representation. Fig. 3. presents the results of comparison of the generating time of the probabilistic-fuzzy knowledge base as the function of the minimal support value for the modified Apriori and FP-Growth algorithms. The chart presents the advantage of modified FP-Growth algorithm.

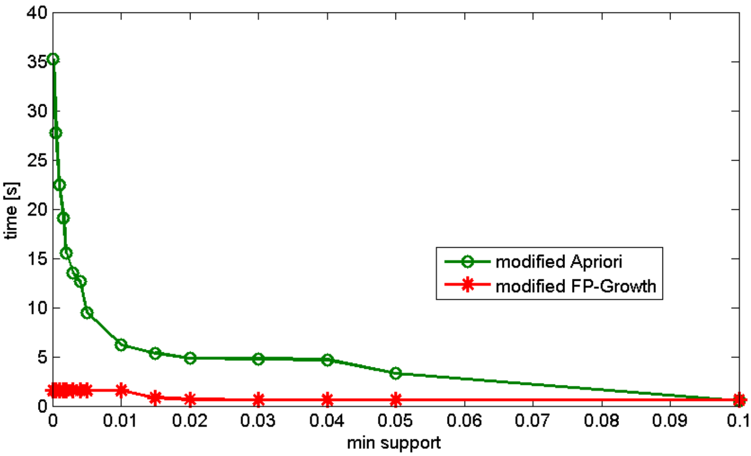

Fig. 3 The time of generating the probabilistic-fuzzy knowledge base as the function of the minimal support value for the modified Apriori and FP-Growth algorithms (3 input variables, one output variable, 5 fuzzy sets for each variable, near 500 learning data)

5 Exemplary Calculations

The created in the work [13] inference system has been applied to predict the values of time varying variables determining the natural phenomena such as wind speed and ash contents (incombustible matter) in row coal. Both variables are very difficult for prediction because of many random rates influence on the measurements results.

The set of 11 000 measurements of the wind speed $X(t)=\{v(t)\}$ t=1,2,…n were recorded at 1-minute samplings. The averages of measurements from 4 steps were researched. First 2000 measurements were treated as learning data, the remaining ones – test data. The forecasts of wind speed $v(t)$ have been made on the grounds of the last three measurements of wind speed denoted as $v(t\text{-}3)$, $v(t\text{-}2)$, $v(t\text{-}1)$. For each variable, in the space of process values, 9 fuzzy sets have been defined, with the linguistic values describing the wind speed, as: "very light", "light", "mild", "moderate", "fairly strong", "strong", "very strong", "squally", "very squally", assuming 45 disjoint intervals of the variables values. Exemplary values of the membership functions for variable $v(t\text{-}3)$ are shown in Fig. 4.

The membership grades for other variables have been analogically defined.

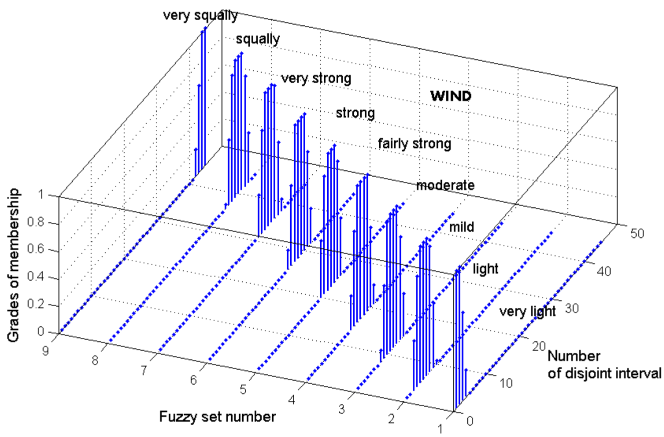

Fig. 4 Fuzzy sets defined for 9 linguistic values of the linguistic variable 'speed wind'

In Table 1. the exemplary joint empirical probability distribution for two chosen linguistic random variables has been presented. We can see that variables take their three from nine linguistic values and the probability distribution is 'narrow', concentrated only over the few linguistic values.

Table 1 Exemplary joint probability distribution of two linguistic random variables (X_t,X_{t-1}) under the conditions: ((X_{t-3} *is moderate) and* (X_{t-2} *is moderate*))

Assumption: wind X(t-3) is 'moderate' and wind X(t-2) is 'moderate'						
X(t-1) \ X(t)	very light	light	mild	moderate	fairly strong	strong
very light	0	0	0	0	0	0
light	0	0	0	0	0	0
mild	0	0	0,0145	0,0196	0	0
moderate	0	0	0,0223	0,0934	0,0180	0
fairly strong	0	0	0	0,0153	0,0156	0
strong	0	0	0	0	0	0

The optimal model structure is derived at the minimal support value, equal to s=0.001, then, the root mean square error for the learning data is 0.5514 m/sec, whereas for the testing data it is 0.6434 m/sec. The model consists of 92 elementary rules (47 file rules). The most important file rules are:

R1: (0.1337) IF ($X(t$-3) IS 'moderate') AND ($X(t$-2) IS 'moderate') AND ($X(t$-1) IS 'moderate')
THEN ($X(t)$ IS 'moderate') (0.6989)
ALSO ($X(t)$ IS 'mild') (0.1665)
ALSO ($X(t)$ IS 'fairly strong') (0.1346)
R2: (0.0973) IF ($X(t$-3) IS 'fairly strong') AND ($X(t$-2) IS 'fairly strong')
AND ($X(t$-1) IS 'fairly strong')
THEN ($X(t)$ IS 'fairly strong') (0.6827)
ALSO ($X(t)$ IS 'moderate') (0.2253)
ALSO ($X(t)$ IS 'strong') (0.0920)
R3: (0.0749) IF ($X(t$-3) IS 'mild') AND ($X(t$-2) IS 'mild') AND ($X(t$-1) IS 'mild')
THEN ($X(t)$ IS 'mild') (0.6683)
ALSO ($X(t)$ IS 'moderate') (0.2131)
ALSO ($X(t)$ IS 'light') (0.1186)

The results of predicted numeric values of the wind speed and measured data have been presented in Fig. 5.

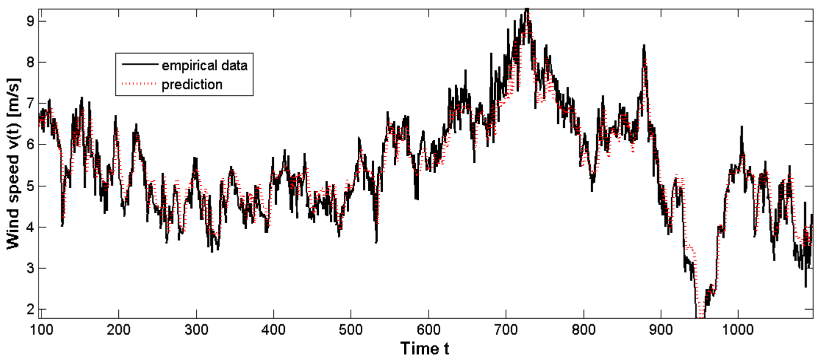

Fig. 5 Comparison of the prediction values and empirical data of wind speed for testing data

The next application of the tested inference system in dynamic system modelling will be shown on the example of coal parameters analysis. Measurements of some coal parameters, as contents of particular density fractions of grains, ash or sulfur contents are the bases of quality control at coal preparation plants and power stations. Sampling research of grain materials are used by technological process engineers to approximate needed coal parameters or characteristics. Sample taking is a random process, according to respective scheme of randomness. In the other hand, experts of technology often express some values of measurements in linguistic categories. These are the reasons that knowledge base of the variation of empirical data concerning coal parameters is created with regards fuzziness and randomness.

The interested reader can find more details on probabilistic-fuzzy modelling characteristics of grain materials in the work [15].

In the example the coegsisting of two variables: the content of light grains fraction, *X*, and the ash content in that fraction, *Y*, in the time series has been analyzed. Spaces of considerations of both variables have been divided into 40 disjoint intervals and 7 triangular fuzzy sets have been defined as the representations of linguistic values: {"very small", "small", "medium small", "medium", "medium large", "large", "very large"}, for both variables.

Derived knowledge base for that dynamic system has a form of 779 file rules. The most important rules are presented below:

R1: (0.0487) IF (*X*(*t*-2) IS 'med. large') AND (*Y*(*t*-2) IS 'small') AND (*X*(*t*-1) IS 'med. large') AND (*Y*(*t*-1) IS 'small') THEN (*Y*(*t*) IS 'small') (0.5277)
ALSO (*Y*(*t*) IS 'v. small') (0.2004)
ALSO (*Y*(*t*) IS 'med. small') (0.1954)
ALSO (*Y*(*t*) IS 'medium') (0.0398)
ALSO (*Y*(*t*) IS 'med. large') (0.0257)
ALSO (*Y*(*t*) IS 'large') (0.0102)
ALSO (*Y*(*t*) IS 'v. large') (0.0009)

R2: (0.0247) IF (*X*(*t*-2) IS 'medium') AND (*Y*(*t*-2) IS 'small') AND (*X*(*t*-1) IS 'med. large') AND (*Y*(*t*-1) IS 'small') THEN (*Y*(*t*) IS 'small') (0.4673)
ALSO (*Y*(*t*) IS 'med. small') (0.2454)
ALSO (*Y*(*t*) IS 'v. small') (0.1478)
ALSO (*Y*(*t*) IS 'medium') (0.1293)
ALSO (*Y*(*t*) IS 'med. large') (0.0075)
ALSO (*Y*(*t*) IS 'large') (0.0018)
ALSO (*Y*(*t*) IS 'v. large') (0.0009)

In Table 2. the probability distribution of two linguistic random variables (Y_t, Y_{t-1}), under the conditions: ((X_{t-2} *is medium large*) *and* (Y_{t-2} *is small) and* (X_{t-1} *is medium large*)), has been presented. It is easy to observe that occurrence of any linguistic value of any variable is possible with a probability grater then zero. This is different distribution then in the first example.

Table 2 Exemplary probability distribution of two linguistic random variables (Y_t, Y_{t-1}) under the conditions: ((X_{t-2} *is medium large*) *and* (Y_{t-2} *is small) and* (X_{t-1} *is medium large*))

Assumption: *X*(*t*-2) is 'med. large' and *Y*(*t*-2) is 'small' and *X*(*t*-1) is 'med. large'							
Y(*t*-1) \ *Y*(*t*)	'v. small'	'small'	'medium small'	'medium'	'medium large'	'large'	'v. large'
'v. small'	0,0047	0,0087	0,0028	0,0002	0,0004	0,0010	0,0002
'small'	0,0097	0,0257	0,0095	0,0019	0,0012	0,0005	4,37E-05
'med. small'	0,0029	0,0063	0,0041	0,0005	0,0002	0,0001	0
'medium'	0,0009	0,0021	0,0008	0,0005	0,0003	9,29E-05	0
'med. large'	0,0003	0,0003	0,0002	0,0001	9,985E-06	0	0
'large'	2,07E-05	0,0005	5,119E-05	0	0	0	0
'v. large'	2,099E-06	9,38E-05	8,362E-05	0	0	0	0

The computed marginal and conditional probability distributions have been used in the prediction procedure. The optimal structure of the model has been derived at the complete probability distribution, by using both Larsen's inference rule and Fodor's t-norm as a representation of the logic AND (see T_1 in chapter 3.4.). Than the root mean square for training data was equal to 0.87, and for testing data 1.85.

6 Conclusions

The use of fuzzy logics in the knowledge-based system makes it possible to express incomplete and uncertain information in a natural language, typical for expression and cognition of human beings. In addition, the application of the probability of events expressed in linguistic categories enables the adjustment of the model on the grounds of numerical information derived from the data stored in the course of the operation of a given real processes. The created model becomes easier for interpretation by its users, what is very important in the strategic decision-making situations, as well as, in diagnostic systems.

References

[1] Agrawal, R., Imielinski, T., Swami, A.: A Mining association rules between sets of items in large databases. In: ACM Sigmod Intern. Conf. on Management of Data, Washington D.C., pp. 207–216 (May 1993)
[2] Box, G.E.P., Jenkins, G.M.: Time Series Analysis; Forecasting and Control. Holden-Day, San Francisco (1976)
[3] Doob, J.L.: Stochastic Processes. Wiley, New York (1953)
[4] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. AI Magazine, 37–54 (1996)

[5] Fisz, M.: Probability and Statistics Theory. PWN, Warsaw (1967) (in Polish)
[6] Frawley, W., Piatetsky-Shapiro, G., Matheus, C.: Knowledge Discovery in Databases: An Overview. AI Magazine, 57–70 (1992)
[7] Han, J., Pei, H., Yin, Y.: Mining Frequent Patterns without Candidate Generation. In: Proc. Conf. Management of Data, SIGMOD 2000, Dallas, TX. ACM Press, New York (2000)
[8] Hellendoorn, H., Driankov, D. (eds.): Fuzzy Model identification, Selected Approaches. Springer, Berlin (1997)
[9] Hüllermeier, E.: Fuzzy methods in machine learning and data mining: status and prospects. Fuzzy Sets and System 156(3), 387–406 (2005)
[10] Łęski, J.: Neuro-Fuzzy Systems. WNT, Warsaw (2008) (in Polish)
[11] Oh, S.-K., Pedrycz, W., Park, K.-J.: Identification of fuzzy systems by means of genetic optimization and data granulation. Journal of Intelligent & Fuzzy Systems 18, 31–41 (2007)
[12] Papoulis, A.: Probability, Random Variables, and Stochastic Processes. WNT, Warsaw (1972) (Polish edition)
[13] Rudnik, K.: Conception and implementation of the inference system with the probabilistic-fuzzy knowledge base. Dissertation, Opole University of Technology, Opole, Poland (2011) (in Polish)
[14] Rudnik, K., Walaszek-Babiszewska, A.: Probabilistic-fuzzy knowledge-based system for managerial applications. Management and Production Engineering Review 3 (in press 2012)
[15] Walaszek-Babiszewska, A.: Fuzzy probability for modelling of particle preparation processes. In: Proc. 4th Int. Conf. Intelligence Processing and Manufacturing of Materials (IPMM 2003), Sendai, Japan (2003)
[16] Walaszek-Babiszewska, A.: IF-THEN Linguistic Fuzzy Model of a Discrete Stochastic System. In: Cader, A., Rutkowski, L., Tadeusiewicz, R., Żurada, J. (eds.) Artificial Intelligence and Soft Computing, pp. 169–174. Academic Publishing House EXIT, Warsaw (2006)
[17] Walaszek-Babiszewska, A., Błaszczyk, K.: A modified Apriori algorithm to generate rules for inference system with probabilistic-fuzzy knowledge base. In: Proc. 7th Int. Workshop on Advanced Control and Diagnosis, November 19-20, CD-ROM, Zielona Góra, Poland (2009)
[18] Walaszek-Babiszewska, A.: Fuzzy Modeling in Stochastic Environment; Theory, knowledge bases, examples. LAP LAMBERT Academic Publishing, Saarbrücken (2011)
[19] Yager, R.R., Filev, D.: Essentials of Fuzzy Modeling and Control. John Wiley and Sons (1994)
[20] Zadeh, L.A.: Probability Measures of Fuzzy Events. Journal of Mathematical Analysis and Applications 23(2), 421–427 (1969)
[21] Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning. Part I. Information Sciences 8, 199–249 (1975); Part II. 8, 301–357, Part III. 9, 43–80