Proceedings of the International Multiconference on
Computer Science and Information Technology, pp. 141 – 150
ISSN 1896-7094

Linguistic Knowledge Representation for Stochastic Systems

Anna Walaszek-Babiszewska

Opole University of Technology,
Institute of Control and Computer Engineering,
Sosnkowski 31, 45-272 Opole, Poland
a.walaszek-babiszewska@po.opole.pl
http://www.po.opole.pl

Abstract. The paper deals with an idea of a linguistic knowledge representation and a linguistic inference. Relational linguistic fuzzy model with weights of rules is utilising. The probability of linguistic values of antecedent and consequent variables, calculated according to Zadeh's definition, is proposed to formulate a linguistic fuzzy model of a stochastic system. The linguistic inference procedures and the exemplary MISO model are presented.

1 Introduction

Linguistic fuzzy knowledge-based systems, linguistic preparation of information and linguistic inference procedures are based on the notion of a linguistic variable [1]. Real knowledge-based systems involve both type of information: numeric and linguistic. Knowledge from human experts is expressed in linguistic categories, and information discovered from experiments is numerical in its character. Recently, there is an increasing interest in using both sources of information for creating models of a high accuracy, flexible, and compatible to the features of real systems. It is necessary to define the interfaces for the clear preparation both types of information [2], by using fuzzy meaning and fuzzy description of a numeric value, defined in the work of Zadeh [1].

Stochastic modelling of real economic, biological or technological systems has been developed by many authors, and there are many stochastic models created under different assumptions and hypotheses. In the same time and in the same environment of imprecision, knowledge of human experts is a base for rational decisions, or risk estimations, expressed by market investors in a form of linguistic variables, sentences or rules.

This paper discusses possibilities of utilising linguistic fuzzy model ling and linguistic inference for stochastic vector process. The idea of a linguistic variable [1] has been used for determining linguistic states of the process. The template modelling [7] has been used for the partition of the input-output space and to determine the sets of linguistic variables. The empirical probability distributions of linguistic variables of the model are calculated according to Zadeh's definition [3]. Considered linguistic fuzzy model can be treated as a fuzzy relational model with weights [4, 5]. Weights of the model represent probabilities of fuzzy events in antecedents and consequences of rules [6].

2 Representations of a Linguistic Variable

Linguistic variable which takes values in a form of words or sentences is used for expressing the linguistic knowledge representation about the behaviours of examined systems.

According to Zadeh [1] (also [8] and [2]), the definition of a linguistic variable can be expressed by the quintuple:

$$< x_{name}, L(X), X, G, M > \tag{1}$$

where

x_{name} – the name of the linguistic variable, e.g., *'temperature'* ;

X – the universe of discourse, the domain of the examined variable, e.g. $X = [-40,40] \in R$;

$L(X)$ – the set of linguistic values (terms) that x_{name} takes, e.g. $L(X)=\{low, middle, high\}$;

G – a syntactic rule for generating the names LX from the term-set $L(X)$;

M – a semantic function that assigns to each linguistic value LX a fuzzy meaning $M(LX)$, as a fuzzy subset of X, e.g.:

$$M(low) = \{(x, \mu_{'low'}(x)), x \in X\} \tag{2}$$

where $\mu_{'low'}(x)$ is a grade of membership of $x \in X$ in $M(low)$.

The membership functions define the family fuzzy subsets in X, for every linguistic value from $L(X)$, e.g.:

$$\mu_{'low'}(x) = \begin{cases} 1, & x \in [-40,-20) \\ \dfrac{-x}{20}, & x \in [-20,0] \end{cases}, \tag{3}$$

$$\mu_{'middle'}(x) = \begin{cases} \dfrac{x+20}{20}, & x \in [-20,0) \\ \dfrac{20-x}{20}, & x \in [0,20] \end{cases}, \tag{4}$$

$$\mu_{'high'}(x) = \begin{cases} \dfrac{x}{20}, & x \in [0,20) \\ 1, & x \in [20,40] \end{cases}. \tag{5}$$

In many applications we use the linguistic vector variable $x^T_{name}=(x^1_{name},\dots,x^p_{name})$, which is expressed by p linguistic variables considered in the universe of discourse, e.g. $\boldsymbol{X} = R^p$; $\boldsymbol{x}_{name}$ is the name of the linguistic vector, e.g. ‘ *state of market* ’, ‘ *technological situation* ’ or ‘ *material characteristics* ’; $\boldsymbol{L}(X)$ is the set of linguistic values, e.g. $\boldsymbol{L}(X)=\{weak, good, very\ good\}$, $\boldsymbol{L}(X)=\boldsymbol{L}(X_1) \times \dots \times \boldsymbol{L}(X_p)$.

In knowledge-based systems we can distinguish two dual interfaces, in which the two dual functions are implemented:

- information processing from the linguistic form into numeric, according to the fuzzy meaning *M*,
- information processing from the numeric form into linguistic, by using the fuzzy description *D*.

The fuzzy description $D(x)$ of a crisp value x is a discrete fuzzy subset on the set of linguistic values $L(X)$, defined as [1], [2]:

$$D(x)=\{(LX,\mu_{D(x)}(LX)),\ LX\in L(X)\} \tag{6}$$

where $\mu_{D(x)}(LX)$ is the degree of membership of the linguistic value LX in a discrete fuzzy set $D(x)$.

The following relationship between the membership functions is fulfilling:

$$\mu_{M(LX)}(x)=\mu_{D(x)}(LX),\quad \forall x\in X,\ \forall LX\in L(X)\,. \tag{7}$$

For example, the fuzzy description of the crisp value x =15, for linguistic variable *temperature* given above, is a fuzzy subset

$$D(15)=\{(low,0),\ (middle,0.25),\ (high,0.75)\} \tag{8}$$

or in another notation

$$D(15)=0/low+0.25/middle+0.75/high\ . \tag{9}$$

3 Stochastic System with Linguistic States

Consider a certain real system, represented by the ordered pair (x,y) of input and output (states) variables:

$$\{(x(t,\omega),y(t,\omega)):\ t\in T,\ x\in X,\ y\in Y,\ \omega\in\Omega\} \tag{10}$$

where X is the system input domain, Y is the output (states) 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. We assume the probabilistic measure *p(x,y)* on a set of realizations of the processes $[x(t), y(t)]$, $t=t_k, k=1,2,\ldots K$ observed at the discrete moments.

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

$$y(t_k)=f[x(t_k), x(t_{k-1}),\ldots, x(t_{k-n}),\ldots, y(t_{k-1}),\ldots, y(t_{k-m})] \quad \textbf{(11)}$$

where *f()* can be a multivariable regression function. These types of models are well known as Box-Jenkins' time series models [9].

Other types of models take into account a multivariable distribution function of the processes $[x(t), y(t)]$, $t=t_k, k=1,2,\ldots K$ observed at the discrete moments

$$p(x, y)=p[x(t_k), x(t_{k-1}),\ldots, x(t_{k-n}),\ldots, y(t_k), y(t_{k-1}),\ldots, y(t_{k-m})] \quad \textbf{(12)}$$

.

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).

Consider now *(x,y)* as a pair of linguistic variables and two types of fuzzy models that can give a fuzzy representation of the stochastic process: Mamdani type fuzzy model (13) and Takagi-Sugeno fuzzy model (15) (according to [10], [5]):

$$\begin{aligned} &w_i[\, IF\ x(t_k)\ is\ A_{i,k}\ AND\ x(t_{k-1})\ is\ A_{i,k-1}\ldots AND\ x(t_{k-n})\ is\ A_{i,k-n}\cdots \\ &AND\ y(t_{k-1})\ is\ B_{i,k-1}\ldots AND\ y(t_{k-m})\ is\ B_{i,k-m} \\ &THEN\ y(t_k)\ is\ B_{i,k}\,] \end{aligned} \quad \textbf{(13)}$$

where *i=1,...,I* – numbers of rules, *x,y* –linguistic variables, $x \in X$, $y \in Y$ with linguistic values sets *L(X), L(Y),* determining linguistic states of the system, $A_{i,k}, A_{i,k-1},\ldots, A_{i,k-n}$ - fuzzy subsets corresponding to linguistic values of variables $x(t_k), x(t_{k-1}),\ldots, x(t_{k-n})$, $x \in X$,

$B_{i,k}, B_{i,k-1},\ldots, B_{i,k-m}$ - fuzzy subsets corresponding to linguistic values of variables $y(t_k), y(t_{k-1}),\ldots, y(t_{k-m})$, $y \in Y$, w_i - weight of the *i-th* rule, probability of the fuzzy event representing the fuzzy relation R_i in $X^{n+1} \times Y^{m+1}$ (according to [6, 11]):

$$w_i=P(R_i)=P(A_{i,k}\times A_{i,k-1}\times\ldots\times A_{i,k-n}\times B_{i,k}\times B_{i,k-1}\times\ldots\times B_{i,k-m}) \quad \textbf{(14)}$$

$$\begin{aligned} &IF\ x(t_k)\ is\ A_{i,k}\ AND\ x(t_{k-1})\ is\ A_{i,k-1}\ldots AND\ x(t_{k-n})\ is\ A_{i,k-n}\cdots \\ &AND\ y(t_{k-1})\ is\ B_{i,k-1}\ldots AND\ y(t_{k-m})\ is\ B_{i,k-m} \\ &THEN\ y_i(t_k)=a_{i,0}x(t_k)+a_{i,1}x(t_{k-1})+\ldots+a_{i,n}x(t_{k-n}) \\ &+b_{i,1}y(t_{k-1})+\ldots+b_{i,m}y(t_{k-m}) \end{aligned} \quad \textbf{(15)}$$

where $a_{i,0}, a_{i,1}, \ldots, a_{i,n}, b_{i,1}, \ldots b_{i,m}$ are the parameters of the regression equation in the *i-th* part of the $X^{n+1} \times Y^{m+1}$ domain, calculated by using e.g. the least square method and data sets.

Mamdani type fuzz y model with weights, as (13), allows to express the fuzzy and probability uncertainty, by using the frequency of fuzzy events. Takagi-Sugeno fuzzy model (15) levels the probabilistic uncertainty by using the regression method. Both models use the information granulation, made by the fuzzy partition of the consideration space.

4 General Form of the MISO Linguistic Fuzzy Model

4.1 Rule Base

A linguistic fuzzy model consists of a collection of rules, the number of rules is a combination of all linguistic values of variables in the antecedent and the consequent, in the form [10, 5]:

$$\begin{array}{l} R_i\colon w_i(IF\ \ x_1 is\ A_{1,i}\ AND\ x_2 is\ A_{2,i} \ldots AND\ x_p is\ A_{p,i} \\ THEN\ \ y\ is\ B_1\ (w_{1/i}) \\ ------------ \\ ALSO\ y\ is\ B_j\ (w_{j/i}) \\ ------------ \\ ALSO\ y\ is\ B_J\ (w_{J/i})) \end{array} \tag{16}$$

where *i=1,...,I* – numbers of rules, $A_{1,i}, \ldots, A_{p,i}$ - linguistic values of the antecedent variables $x^T = [x_1 \ldots x_p],\ x \in R^p$, defined by corresponding membership functions $\mu_{A_{1,i}}(x_1), \ldots \mu_{A_{p,i}}(x_p)$, $A_{1,i} \in L(X_1), \ldots, A_{p,i} \in L(X_p)$; B_j, *j=1,...,J* - linguistic values of the consequent variable $y \in Y \subset R$, defined by corresponding membership functions $\mu_{B_j}(y)$, $B_j \in L(Y)$;

w_i, $w_{j/i}$, *i=1,...,I, j=1,...,J* - weights of rules, representing probabilities of fuzzy events in antecedents and consequents of the model, calculated according to Zadeh's definition [3], as follows (according to [6] and [11]):

w_i is a marginal probability of a joint fuzzy event in the antecedent

w_{ij} is a joint probability of the fuzzy event (fuzzy relation) in the antecedent and consequent of the elementary rule

$$w_{ij} = P(A_i \times B_j) = \sum_{(x,y) \in X \times Y} p(x,y) T\left(\mu_{A_i}(x), \mu_{B_j}(y)\right) \quad , \tag{18}$$

$w_{j/i}$ is a conditional probability of the consequent fuzzy event

$$w_{j/i} = P(B_j / A_i) = \frac{P(A_i \times B_j)}{P(A_i)}, \tag{19}$$

and *p(x,y)* is a joint probability function, in a classical sense, which assigns to each Borel's set in $X \times Y$ a real number $p \in [0,1]$.

The weights fulfil the relations

$$\sum_{i=1,2,\ldots,I} w_i = 1, \quad \sum_{i=1,\ldots,I} \sum_{j=1,\ldots,J} w_{ij} = 1, \quad \sum_{j=1,\ldots,J} w_{j/i} = 1. \tag{20}$$

The model (16) can be also written as:

$$R_i : w_i (IF\ x_1\ is\ A_{1,i}\ AND\ \ldots AND\ x_p\ is\ A_{p,i}\ THEN\ y\ is \sum_{B_j \in L(Y)} w_{j/i} / B_j) \tag{21}$$

4.2 Linguistic Inference

The linguistic fuzzy model (16) describes the fuzzy relation *R* on the linguistic space $L(X_1) \times \ldots \times L(X_p) \times L(Y)$. According to linguistic inference presented in [2] and using the notion of a probability of linguistic values of the model variables [11], the relation can be written as:

$$\forall A_{1,i} \in L(X_1), \ldots, \forall A_{p,i} \in L(X_p), \\ \forall B_j \in L(Y) : \mu_R(A_{1,i}, \ldots, A_{p,i}, B_j) = w_{ij} \tag{22}$$

The linguistic fuzzy output

$$D(y) = \left\{ \left(B_j, \mu_{D(y)}(B_j) \right), B_j \in L(Y) \right\}, j = 1, \ldots, J \tag{23}$$

can be inferred by using Zadeh's compositional rule and fuzzy input defined on $L(X_1) \times \ldots \times L(X_p)$. The membership function for every linguistic value $B_j,\ j = 1, \ldots, J$ can be given by the expression:

$$\mu_{D(y)}(B_j) = \underset{i}{S} \left(T(T(\mu_{D(x_1)}(A_{1,i}), \ldots, \mu_{D(x_p)}(A_{p,i})), \mu_R(A_{1,i}, \ldots, A_{p,i}, B_j)) \right) \tag{24}$$

where T, S are a t-norm and t-conorm, respectively.

5 Exemp lary Linguistic Fuzzy Model of the Stochastic Process

The set of 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 will be modelled to predict some linguistic value of the process and the probability of its occurrence.

We assume two variables with delays: $x_{t_{k-1}}, x_{t_{k-2}}$ as antecedent variables, so the created model will represent a fuzzy relation $R(x_{t_k}, x_{t_{k-1}}, x_{t_{k-2}})$ of the linguistic variables in a form of rules (16). Three linguistic states of the process have been distinguish: *L(X)={low, middle, high}* and the fuzzy meaning have been defined in the process domain X:

‘ *low value of x* ’: $A_1 = 1/a_1 + 0.5/a_2$,

‘ *middle value of x* ’: $A_2 = 0.5/a_2 + 0.5/a_3$,

‘ *high value of x* ’: $A_3 = 0.5/a_3 + 1/a_4$,

where $a_1, \ldots, a_4$ are disjoint intervals in $X = [3.4,\ 4.2] \in R$.

T he 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,…,4.

The 3D empirical probability distribution $P(x_{t_{k-1}}, x_{t_{k-1}}, x_{t_{k-2}})$ of linguistic variables, taking the fuzzy states A_1, A_2, A_3 observed at moments t_k, t_{k-1}, t_{k-2} has been computed, according to (18). Then, the marginal and conditional probability distributions of fuzzy states have been calculated, by summation and using (19), respectively.

The ling uistic fuzzy model consists of 7 file rules, because two weights representing probabilities of fuzzy events in antecedents are equal to zero: $P((x_{t_{k-2}} is\ low) \cap (x_{t_{k-1}} is\ high)) = 0$,

$P((x_{t_{k-2}} is\ hihg) \cap (x_{t_{k-1}} is\ low)) = 0$.

The first rule is as follows:

$$R_1 : 0.365(IF\ x_{t_{k-2}} is\ low\ AND\ x_{t_{k-1}} is\ low \quad THEN\ y\ is\ 0.84/low + 0.16/middle\ ) \tag{25}$$

Table 1. presents all 7 file rules of the created model in a form of a decision table. The weights (real numbers) represent values of conditional and marginal probabilities, as in (21) and (25). It is easy to note, that $\sum_{i=1,\ldots,7} w_i = 1$ and $\sum_{j=1,2,3} w_{j/i} = 1,\ i = 1,\ldots,7$.

We show the linguistic inference algorithm taking two exemplary numeric values of the process: $x_{t_{k-2}} = 39.5,\ x_{t_{k-1}} = 40.5$ and the elaborated fuzzy model. We want to predict the linguistic fuzzy set B of the consequent variable x_{t_k}

$$B = \mu_B(low)/low + \mu_B(middle)/middle + \mu_B(high)/high \tag{26}$$

Table 1. The decision table representing the rule-based fuzzy model of the examined stochastic process

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

By applying the fuzzy description notion (6) and the assumed partition of the input-output space, we find linguistic fuzzy sets of the antecedent variables, as follows:

$$D_{(39.5)}(x_{t_{k-2}}) = 0.5/middle + 0.5/high$$

$$D_{(40.5)}(x_{t_{k-1}}) = 1/high\,.$$

The membership functions for estimated linguistic fuzzy set B can be calculated according to (24), as follows:

$$\mu_B(low) = S\{T\big(T\big(\mu_{D(39.5)}(middle), \mu_{D(40.5)}(high)\big), \mu_R(middle, high, low)\big), \tag{27}$$
$$T\big(T\big(\mu_{D(39.5)}(high), \mu_{D(40.5)}(high)\big), \mu_R(high, high, low)\big)\}$$

The value of the expression (27) is equal to zero, because the membership values of the relations are equal to zero: $\mu_R(middle, high, low) = 0$, $\mu_R(high, high, high) = 0$
(see Table 1.) and t-norms are also equal to zero.

$$\mu_B(midle) = S\{T(T(\mu_{D(39.5)}(middle), \mu_{D(40.5)}(high)), \mu_{R_6}(midle, high, middle)), T(T(\mu_{D(39.5)}(high), \mu_{D(40.5)}(high)), \mu_{R_7}(high, high, middle))\} \quad (28)$$

$$\mu_B(high) = S\{T(T(\mu_{D(39.5)}(middle), \mu_{D(40.5)}(high)), \mu_{R_6}(middle, high, high)), T(T(\mu_{D(39.5)}(high), \mu_{D(40.5)}(high)), \mu_{R_7}(high, high, high))\} \quad (29)$$

Applying a product t-norm $T(a,b)=ab$ and a bounded sum $S(a,b)=\min(1, a+b)$ as a t-conorm to the expressions (28) and (29) we calculate the fuzzy linguistic value *B* of the consequent variable x_{t_k}, as follows:

$$\mu_B(midle)=S\{T(T(0.5, 1), 0.48), T(T(0.5, 1), 0.7)\}= S(0.24, 0.035)=0.275 \quad (30)$$

$$\mu_B(high)=S\{T(T(0.5, 1), 0.52), T(T(0.5, 1), 0.93)\}= S(0.26, 0.465)=0.725 \quad (31)$$

Output discrete fuzzy set, determined on the linguistic term set *L(X)={low, middle, high}*, representing the forecast of the linguistic variable *x* and the levels of affiliation to linguistic values is equal to:

$$B=0/low + 0.275/middle + 0.725/high \quad (32)$$

We can also use one of the defuzzyfication methods to calculate the numeric value of $x^*_{t_k}$ [2]. Both, the fuzzy linguistic set (32) and a possible numeric value state the predicted values of the linguistic consequent variable. We can also use the probability of events $P((x_{t_{k-2}} isA_i) \cap (x_{t-1} isA_j))$ to compute the expected value of predicted numeric values. Details are shown in [11].

References

1. Zadeh L. A.: *The Concept of a Linguistic Variable and its Application to Approximate Reasoning,* Information Sciences, Part I: 8 (1975) 199-249.
2. Nakoula S., Galichet S., Foulloy L.: *Identification of Linguistic Fuzzy Models Based on Learning,* In: Hellendoorn H., Driankov D. (eds): Fuzzy Model Identification; Selected Approaches. Springer-Verlag, Berlin Heidelberg (1997) 281-319.
3. Zadeh L. A.: *Probability Measures of Fuzzy Events,* Journal of Mathematical Analysis and Applications, Vol. 23, No 2 (1968) 421-427.
4. Pedrycz W.: *An Identification Algorithm in Fuzzy Relational Systems,* Fuzzy Sets and Systems, Vol. 13, (1984) 153-167.

5. Rutkowski L.: *Methods and Techniques of Artificial Intelligence* (in Polish). PWN, Warsaw (2005).
6. Walaszek-Babiszewska A.: *IF-THEN Linguistic Fuzzy Model of a Discrete Stochastic System,* In: Cader A., Rutkowski L., Tadeusiewicz R., Żurada J. (eds): Book series: Challenging Problems of Sciences; Computer Sciences. EXIT, Warsaw (2006) 169-174.
7. Yager R. R., Filev D. P.: *Template Based Fuzzy Systems Modelling. Journal of Intelligent Systems,* Vol. 2 (1994) 39-54.
8. Kacprzyk J.: *Fuzzy Sets in System Analysis* (in Polish). PWN, Warsaw (1986).
9. Box G. E. P., Jenkins G. M.: Time Series Analysis; Forecasting and Control (in Polish). PWN, Warsaw (1983).
10. Yager R. R., Filev D. P.: *Essentials of Fuzzy Modeling and Control* (in Polish). WNT and Wiley, Warsaw (1995).
11. Walaszek-Babiszewska A.: *Construction of Fuzzy Models Using Probability Measures of Fuzzy Events,* In: Kaszyński R. (ed): Proc. 13 th IEEE/IFAC International Conf. 'Methods and Models in Automation and Robotics', CD, Szczecin, Poland (2007) 661-666.