On linguistic representation of quantitative dependencies

Ildar Batyrshin\textsuperscript{a,b}

\textsuperscript{a}Instituto Mexicano del Petroleo, Mexico City 07730, Mexico
\textsuperscript{b}Institute of Problems of Informatics, Academy of Sciences of Tatarstan, Tatarstan, Russian Federation

Abstract

A description of quantitative dependencies by a novel type of fuzzy rules like ‘If $X$ is SMALL then $Y$ is QUICKLY INCREASING’ is considered. The use of such rules for representation of perception based and numerical information about dependencies between variables is discussed. These rules are based on a granulation of directions of function change or slope values. Perception based information given by rules is represented by a fuzzy function $Y(X)$. A method of solution of a fuzzy equation $Y(X) = B$ is considered. A linguistic representation of given numerical information about dependencies between variables $X$ and $Y$ is based on a fuzzy partition of the domain of $X$ on fuzzy intervals, on a linear approximations of data on these intervals and on a linguistic retranslation of results. A genetic algorithm is used for obtaining fuzzy partitions. In conclusion, the possible applications of proposed methods in petroleum industry are discussed.

Keywords: Fuzzy relation; Granular model; Soft computing; Computing with words

1. Introduction

Fuzzy logic based computing methodologies give regular ways for dealing with linguistic and numerical information encountered in real-world problems (Babuska, 1998; Jang, Sun, & Mizutani, 1997; Kosko, 1997; Nikravesh & Aminzadeh, 2001; Wang, 1997; Zadeh, 1966; Zadeh, 1997; Zadeh, 1999). Integration of linguistic and numerical information is important for insertion of expert knowledge in formal models, for interpretation of numerical data, for description and formalization of uncertain information. The integration in the framework of soft computing of several methodologies based on fuzzy logic, neural networks, genetic algorithms, etc. gives possibility to create hybrid intelligent systems for solution of problems in the presence of imprecise, incomplete and uncertain information. Such systems find wide applications in pattern recognition, forecasting, decision-making and control systems.

The most popular fuzzy models used in soft computing are based on the rules of two types (Jang et al., 1997; Kosko, 1997; Wang, 1997). Mamdani models in the simplest form consist of the rules like the following:

\begin{align*}
R_1 & : \text{If TEMPERATURE is LOW then DENSITY is MIDDLE}, \\
R_2 & : \text{If TEMPERATURE is HIGH then DENSITY is VERY SMALL},
\end{align*}

where LOW, HIGH, VERY SMALL and MIDDLE are fuzzy sets defined on the domains of numerical values of temperature $t$ and density $d$. The suitable definition of membership functions, logical operations and inference procedure determine the correspondence between the set of rules $\{R_i\}$ and real functions $d = f(t)$.

Takagi-Sugeno-Kang (TSK) fuzzy models differ from Mamdani models due to the presence of functions in the consequent parts of rules.

\begin{align*}
R_1 & : \text{If TEMPERATURE is LOW then } d = f_1(t), \\
R_2 & : \text{If TEMPERATURE is HIGH then } d = f_2(t),
\end{align*}

where $f_i$ are real (usually linear) functions of real variable $t$. TSK model also determines some real function $d = f(t)$ which is obtained as weighted sum of values obtained in antecedent parts of rules.

Mamdani and TSK models are universal approximators of functions but usually TSK models due to the presence of functions in the consequent parts of rules require the less number of rules and they are easily tuned. From another point of view the completely linguistic form of rules in Mamdani models have advantages in the representation of expert knowledge and in the linguistic interpretation of dependencies.

Based on these considerations the new type of rules combining the properties of rules in Mamdani and TSK models have been introduced by Batyrshin (2002) and Batyrshin and Panova (2001). These rules like rules in

E-mail address: batyr1@hotbox.ru (I. Batyrshin).
Mamdani models have completely linguistic form and like rules in TSK models contain the function description in the consequent part of rules. Below are the examples of such rules:

\[ R_1 : \text{If TEMPERATURE is LOW then DENSITY is SLOWLY INCREASING,} \]

\[ R_2 : \text{If TEMPERATURE is HIGH then DENSITY is QUICKLY DECREASING.} \]

The rules like (1) and (2) may reflect the human perceptions about influence of one parameter on another. These perceptions may represent a compressed generalized description of statistical and functional dependencies obtained as a result of series of measurements. These perceptions may be based also on a human experience which does not accompanied by results of measurements.

Very often the dependencies \( z = f(x, y) \) between parameters of system are represented by the family of curves \( z_i = f_i(x) \) for some limited number of values of parameter \( y \). Usually it arises from absence of necessary information due to the difficulties in calculation or measurement of dependencies \( z_i = f_i(x) \) for a large number of parameter values \( y \). In this case the linguistic description of the curves \( z_i = f_i(x) \) like in rules (1) and (2) may be used for a generalized description of the curves in some interval of values of parameter \( y \). For example, the rules (1) and (2) may describe dependence \( D(T) \) in some chemical reactor for the values of PRESSURE greater than some value \( P^* \) but for small values of PRESSURE the functional dependence between \( D \) and \( T \) may be described by another set of rules, e.g.:

\[ R_1 : \text{If TEMPERATURE is LOW then DENSITY is INCREASING,} \]

\[ R_2 : \text{If TEMPERATURE is HIGH then DENSITY is VERY QUICKLY DECREASING.} \]

The goal of a numerical modeling of processes very often consists in the search of the values of parameters of processes when the functional dependencies between parameters are qualitatively changing (Fletcher, 1988; Gilmanov & Panova, 1999). For this reason the development of methods of processing of new type of rules may be used for qualitative reasoning about systems (De Kleer & Brawn, 1984; Forbus, 1984; Kuipers, 1984).

Another example gives time series of market prices on agricultural goods. Such time series usually change from year to year but the seasonal changes have usually the similar tendencies. The human perceptions about such tendencies may be also represented in linguistic form like

\[ R_1 : \text{In JUNE THE PRICE ON SUGAR is INCREASING,} \]

\[ R_2 : \text{In THE END OF OCTOBER THE PRICE ON SUGAR is QUICKLY DECREASING,} \]

where JUNE and THE END OF OCTOBER are fuzzy intervals. It should be noted that in such cases the human perceptions are based on some statistics but the statistical data are usually not written but conserved as human perceptions.

Two different approaches to construction of such rules are discussed in the following sections. The first is based on a formalization of perception based expert knowledge given linguistically (Batyrishin, 2002; Batyrshin & Panova, 2001). The set of such rules may be considered as a linguistic description of derivative of function and a reconstruction of function may be considered as a solution of initial value problem based on such linguistic derivative. A reconstruction of function is based on operations of projection, cylindrical extension in direction, etc. introduced by Zadeh (Zadeh, 1966, 1997). The solution of linguistic initial value problem may be considered as a model of computing with words (Zadeh, 1999). In contrast to Batyrshin (2002), we consider here a method of solution of a fuzzy equation \( Y(X) = B \) formulated linguistically: Find the value of \( X' \) when the function \( Y = Y(X) \) will be equal to \( Y' = B \). Such problem arises when it needs to obtain a desired output of the system when the dependencies between the parameters of a system are given linguistically by rules like (1) and (2).

In conclusion we briefly discuss the possibility of application of such approach to construction of intelligent training system based on expert model of chemical reactor in petrochemical industry.

The second approach is based on a linguistic description of given numerical dependencies (Batyrishin & Wagenknecht, 2002). This approach is based on a partition of the domain of independent variable on fuzzy intervals, on approximation of data on these intervals and on linguistic interpretation of intervals and approximating lines. In some sense the obtained model may be considered as alternative to TSK model and we discuss the possibility of application of this model instead of TSK model for approximation of data in petroleum industry.

2. Granular differentials and differential equations

Differential equations play important role in mathematical modeling. But often the values of variables used in considered problem are uncertain; moreover the functional dependencies between variables may be unknown. In the first case the model of the process may be based on fuzzy differential equations, i.e. on differential equations with fuzzy parameters (Nieto, 1999; Vorobiev & Seikkala, 2002). In the second case the model of the process may be based on a qualitative description which uses instead of derivatives the signs of derivatives or, equivalently, the labels ‘increasing’, ‘steady’ and ‘decreasing’ (De Kleer & Brawn, 1984; Forbus, 1984; Kuipers, 1984). The granular rule-based approach to representation of derivatives
occupies the intermediate position between these two approaches.

The rules like (1) and (2) are considered as linguistic expressions of dependencies between variables $Y = DENSITY$ and $X = TEMPERATURE$, such that $Y$ is $SLOWLY INCREASING$ function of $X$ on fuzzy interval $LOW$ and $Y$ is $QUICKLY DECREASING$ function of $X$ on fuzzy interval $HIGH$. The linguistic label $SLOWLY INCREASING$ in rule (1) may be interpreted as a linguistic evaluation of the speed of the change of variable $Y = DENSITY$ when the variable $X = TEMPERATURE$ is increasing within the fuzzy interval $LOW$. Since the speed of function change is related with the derivative of the function, the right side of this rule may be considered also as a linguistic evaluation of derivative value $dY/dX$ on this interval. In terms of derivatives, the rules (1) and (2) may be translated in the following form:

$R_1$: If $X$ is $LOW$ then $dY/dX$ is $POSITIVE SMALL$. (3)

$R_2$: If $X$ is $HIGH$ then $dY/dX$ is $NEGATIVE LARGE$. (4)

Since the value of derivative equals to the slope of the tangent line to the curve of function, the linguistic labels in right sides of rules may be considered also as linguistic evaluations of this slope or evaluations of parameter $p$ in the equation of tangent line $\gamma = px + q$. A direction of the change of the function defined by the tangent will be represented by a fuzzy granule of directions. From another point of view the granule of directions defines fuzzy sets of differential values $dY$ correspondent to given crisp values of increment $\Delta x$ as follows: $dY = P \Delta x$, where $P$ is a granular slope value defined by a rule. We will suppose that the range of crisp values of increment $\Delta x$ is defined by the left part of correspondent rule. As a result, the granular differential $dY$ may be considered as a fuzzy function of crisp argument $\Delta x$. For example, the rule (4) defines a fuzzy differential as a fuzzy function $dY = P \Delta x$, where $P$ is a fuzzy set correspondent to term $NEGATIVE LARGE$ and $\Delta x$ takes values in fuzzy interval defined by the term $HIGH$. For explicitation of rules it is necessary to define a linguistic scales for linguistic variables used in rules, to define a granulation of possible slope values and to establish a correspondence between the grades of scales and fuzzy sets.

Suppose the domain of slope values equals to $[-10, 10]$ and granules of slopes are defined by some fuzzy sets with central modal values $p_i$ $(i = 1, \ldots, 7)$. The example of linguistic scales and centers of membership functions correspondent to linguistic grades of the scales are shown in Table 1. Each grade of the scale represents some fuzzy granule of directions $l_i$.

Methods of construction of granular directions were considered in Batyrshin (2002) and Batyrshin and Panova (2001). We consider here a cylindrical extension in direction and correspondingly cylindrical differential

$$
\mu_{d_i}^c(\Delta x, dy) = \mu_{dY}(dy),
$$

where $\mu_{d_i}^c(dy)$ is a membership function on $dY$ correspondent to direction $l_i$. For example, the cylindrical extension of generalized bell membership functions (GBMF) (Jang et al., 1997) for each value $\Delta x > 0$ may be defined as follows

$$
\mu_{d_i}(dy) = \frac{1}{1 + \left| \frac{dy - dy_i}{a_i} \right|^b},
$$

where $dy_i = p_i \Delta x$, and $a_i$, $b_i$ are parameters of GMBF. For $\Delta x = 0$ we define $D_0^c$ by Eq. (6) with $dy = 0$. $D_0^c$ is called a starting set for cylindrical extension of direction $l_i$. The fuzzy value of cylindrical differential will have constant cross-section. The example of cylindrical differential constructed by GMBF is shown in Fig. 1.

The total set of rules like (3) and (4) with granular derivatives in the right parts of rules may be considered as a granular description of derivative $dY/dX = F(X)$ of function $Y$ piecewise defined on the domain of variable $X$. Each rule defines some piece of derivative on the fuzzy interval correspondent to the value of $X$ in the left part of rule.

The set of terms of linguistic variable $X$ can include the labels $VERY SMALL$, $SMALL$, $MIDDLE$, $LARGE$, $VERY LARGE$, $APPROXIMATELY N$, $BETWEEN N AND M$, $GREATER THAN N$, etc., where $N$ and $M$ are some real values or fuzzy numbers. The meaning of these terms may be explicitated by definition of correspondent fuzzy sets defined on $X$. The set of terms and correspondent fuzzy sets

Table 1

<table>
<thead>
<tr>
<th>$l_i$</th>
<th>Linguistic description of the speed of function change</th>
<th>Linguistic value of derivative (slope)</th>
<th>$p_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>QUICKLY INCREASING</td>
<td>POSITIVE LARGE</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>INCREASING</td>
<td>POSITIVE MIDDLE</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>SLOWLY INCREASING</td>
<td>POSITIVE SMALL</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>CONSTANT</td>
<td>ZERO</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>SLOWLY DECREASING</td>
<td>NEGATIVE SMALL</td>
<td>-3</td>
</tr>
<tr>
<td>2</td>
<td>DECREASING</td>
<td>NEGATIVE MIDDLE</td>
<td>-6</td>
</tr>
<tr>
<td>1</td>
<td>QUICKLY DECREASING</td>
<td>NEGATIVE LARGE</td>
<td>-9</td>
</tr>
</tbody>
</table>

Fig. 1. Cylindrical differentials in directions 3 (‘SLOWLY DECREASING’) and 4 (‘CONSTANT’) based on generalized bell membership functions.
are determined by a granulation of the domain of X. Generally, for the same rule base this granulation may depend on some parameter or context. The granulation of slopes may also depend on the value of some parameter. In this case, the rule base may describe the parametric family of granular derivatives with explicit dependence on the value of the second parameter. For example, the rules (3) and (4) may describe the derivative \( \frac{d}{dX} \text{PRESSURE} \) for different values of the third parameter \( Z \) = \text{PRESSURE} but the membership functions of fuzzy sets correspondent to linguistic values of antecedents and consequents of rules may be different and depend on the value of \( Z \).

Suppose a granular ordinary differential equation
\[
dY/dX = F(X)
\] is given by the rule base
\[
R_i : \text{If } X = A_i \text{ then } dY/dX = P_i, \ (i = 1, ..., m). \tag{8}
\]

The method of solution of such equation satisfying the initial condition:
\[
\text{If } X = X_0 \text{ then } Y = Y_0
\]

where \( X_0, Y_0 \) are fuzzy sets defined on \( X \) and \( Y \), respectively, was proposed in Batyrshin (2002). This method may be considered as a granular generalization of Euler method of sequential propagation of fuzzy set considered as a granular generalization of Euler method of was proposed in Batyrshin (2002). This method may be considered as a reconstruction of perception based function given by the granular differential equation:
\[
R_1 : \text{If } X \text{ is SMALL then } Y \text{ is QUICKLY INCREASING},
\]
\[
R_2 : \text{If } X \text{ is MEDIUM then } Y \text{ is INCREASING},
\]
\[
R_3 : \text{If } X \text{ is LARGE then } Y \text{ is SLOWLY INCREASING},
\]

and by the initial value:
\[
R_4 : \text{If } X \text{ is APPROXIMATELY 0 then } Y \text{ is APPROXIMATELY 10}.
\]

The solution of granular initial value problem may be considered as a reconstruction of perception based function given by rules.

The reply on a query ‘For what value of \( X \) a value of \( Y \) equals to \( Y' = B \)’ may be considered as a solution of fuzzy relational equation \( Y(X) = B \). It may be obtained as a result of execution of the following steps: (1) Calculate a cylindrical extension of the fuzzy set \( B \) along the axis \( X : C_B(B)(x, y) = B(y) \); (2) Intersect \( C_B(B) \) with the fuzzy relation \( R : R(Y')(x, y) = (C_B(B) \cap R(x, y) = \min(B(y), R(x, y))) \); (3) Obtain a result \( X' = A \) by projection of \( R(Y') \) along the axis \( Y : A(x) = \text{Proj}_y(R(Y'))(x) = \max_x(R(Y')(x, y)) = \max_x(\min(B(y), R(x, y))) \). These steps are shown in Fig. 3. Relational operations of cylindrical extension and projection are introduced by Zadeh (1966) and may be found in Jang et al. (1997).

The procedure of linguistic retranslation of the reply may be based on known linguistic approximation procedures (Batyrshin & Wagenknecht, 2002). If it is necessary, a defuzzification procedure may be used for transformation of the fuzzy reply in a numerical value (Jang et al., 1997). The fuzzy reply ‘\( X = \text{BETWEEN 4 AND 6.5} \)’ in Fig. 3d was obtained as follows. A level set \( A_a = \{ x | A(x) \geq a \} \) for \( a = 0.9 \) was replaced by the nearest interval \([4, 6.5]\) such that the borders of interval coincide with some knots of the grid defined on \( X \). As a defuzzified value \( x' = 4 \) it was used the left border of the obtained interval. In practical applications the value of the level \( a \), the value of grid on \( X \) and defuzzification procedure may be considered as parameters of the model. The possible application of perception based functions is discussed in conclusion.

3. Linguistic description of numerical data

The problem of linguistic description of dependencies in data arises in such areas as trend analysis, process monitoring, diagnosis and control, data mining, qualitative reasoning about processes, etc. (Babuska, 1998; Jang et al., 1997; Kivikunnas, 1999; Kosko, 1997; Wang, 1997; Zadeh, 1997, 1999). Here we discuss the problem of linguistic description of dependencies in given data \( \{ x_i, y_i \}, (i = 1, ..., n) \) by a set of rules
\[
R_k : \text{If } X \text{ is } T_k \text{ then } Y = S_k, \tag{9}
\]

where \( T_k \) are linguistic terms like \text{SMALL}, \text{LARGE}, \text{BETWEEN 5 AND 7} describing some fuzzy intervals \( A_k \) on \( X \) and \( S_k \) are linguistic terms like \text{DECREASING} and \text{QUICKLY INCREASING} that characterize the speed of a change of \( y \) on these intervals.

The methods of generation of such set of fuzzy rules proposed in Batyrshin and Wagenknecht (2002) are briefly considered below. The solution of this problem is obtained as a result of approximation of given data by linear functions \( y_k = p_k x + q_k \) on fuzzy intervals \( A_k \) which form fuzzy partition of the domain of \( X \) and by subsequent retranslation of these intervals and slopes \( p_k \) into linguistic terms \( T_k \) and \( S_k \) correspondingly. A genetic algorithm was used for obtaining an optimal fuzzy partition with approximating functions calculated analytically.

The considered problem is divided into two mutually related parts: (1) the construction of optimal ‘admissible’ fuzzy partition of the domain of input variable \( X \) on fuzzy intervals with optimal linear approximation of given data on these intervals, and (2) the linguistic interpretation of the obtained fuzzy intervals and lines. The first problem is connected with the problem of fuzzy piecewise linear approximation of data, approximation by polygons and splines, shape analysis, fuzzy trend analysis, regression
analysis, approximation of data by TSK fuzzy models, etc. (Babuska, 1998; Conte & de Boor, 1972; Finol, Guo & Jing, 2001; Friedman, 1991; Jang et al., 1997; Kivikunnas, 1999; Loncaric, 1998; Mallon & Swarbrick, 2002).

The second problem is related with the theory of fuzzy information granulation and computing with words (Zadeh, 1997, 1999). The proposed approach to acquisition of rule bases from given data may also be used for

Fig. 2. The steps of reconstruction of perception based function: (a) The fuzzy constraints on the variable \( X \) given in the antecedent parts of rules; (b)–(d) The propagation of initial value \( Y_0 \) in the directions given by the consequent parts of rules and restriction of this directions by constraints given in the correspondent antecedent parts of rules; (e) The overall fuzzy function obtained as an aggregation of constrained directions.
calculation of granular derivatives discussed in the previous section.

The solution to the first problem is based on a stepwise application of genetic algorithm together with merging procedure. Initially a partition with sufficiently large number of fuzzy sets is chosen. Genetic algorithm obtains the optimal partition on given number of fuzzy sets and merging procedure select and merge some two classes if the obtained partition does not admissible. A partition is considered as admissible if it satisfies to some requirements. The merging procedure merges the following classes with a neighboring class: (1) the small fuzzy intervals, which are not meaningful for linguistic interpretation; (2) the intervals with small number of given points because small errors in initial data in such intervals may cause large deviations in slopes; (3) the neighboring intervals with similar linguistic interpretation of slopes. After each merging of some neighboring intervals the genetic algorithm is applying again for finding a new optimal partition. The genetic algorithm and the merging procedure are sequentially applied to reduced number of classes until some final optimal admissible partition is found.

Suppose \( D = \{(x_i, y_i)\}, x_i \in X, y_i \in Y, (i = 1, \ldots, n) \) to be a given set of data where \( X \) and \( Y \) are real intervals: \( X = [x_l, x_r], x_l < x_r, Y = [y_l, y_r], y_l < y_r \). Denote \( X_D = \{x_i\}, (i = 1, \ldots, n) \) and \( Y_D = \{y_i\}, (i = 1, \ldots, n) \). The goal is to represent the underlying crisp dependence between input \( x \) and output \( y \) by a set of rule (9).

Denote \( F(X) \) a set of all fuzzy subsets of \( X \) and \( L_X = \{T_1, T_2, \ldots\} \) a set of linguistic terms used for describing elements of \( F(X) \). Suppose a re-translation function \( N_Z : F(X) \rightarrow L_Z \) is defined (Batyrshev & Wagenknecht, 2002). For example, the set \( L_{\text{TEMPERATURE}} = \{\text{SMALL}, \text{ LARGE}, \text{ 100°}, \text{ BETWEEN 30° and 40°}, \text{ NORMAL}, \text{ VERY HOT}, \lambda \} \) may define the values of linguistic variable \( \text{TEMPERATURE} \) where \( \lambda \) denotes a term ‘meaningless’.

Denote \( Z = [p_l, p_r] \) a set of possible slope values \( p \) for functions \( y = px + q \) and \( L_Z = \{S_1, \ldots, S_t\} \) a set of linguistic terms used for notation of slopes like in Table 1. Suppose a re-translation function \( N_Z : F(Z) \rightarrow L_Z \) is given and it translates into linguistic terms the numerical values of slopes considered as singleton fuzzy sets (Batyrshev & Wagenknecht, 2002; Jang et al., 1997).

A function \( y_k = px + q \) is called a linear approximation of \( D = \{(x_i, y_i)\}, x_i \in X, y_i \in Y, (i = 1, \ldots, n) \) on fuzzy set \( A_k : X \rightarrow [0, 1] \) if it minimizes the function

\[
Q(p, q) = \sum_{i=1}^{n} [y_i - y_k(x_i)]^2 A(x_i),
\]

for all possible values of \( p \) and \( q \). The parameters \( p, q \) may be obtained analytically similarly to the solution of least-square approximation problem (Batyrshev & Wagenknecht, 2002; Conte & de Boor, 1972).

Since for given fuzzy partition \( P = \{A_k\}, (k = 1, \ldots, m) \) of \( X \) the approximation functions \( y_k = px_k + q_k \) on \( A_k \) may be calculated analytically and the linguistic retranslations of intervals \( T_k = N_X(A_k) \) and slopes \( S_k = N_Z(p_k) \) may be found then the problem of linguistic description of data \( D \) is reduced to the problem of finding an optimal admissible partition of given set \( X \) with respect to some criteria of optimality.

The problem was formulated as follows (Batyrshev & Wagenknecht, 2002): Find an admissible fuzzy partition \( \{A_k\}, (k = 1, \ldots, m) \), of \( X \) that minimizes the fitness function

\[
Q_t = \sum_{k=1}^{m} \sum_{i=1}^{n} \frac{[y_i - y_k(x_i)]^2 A_k(x_i)}{\sum_{i=1}^{n} A_k(x_i)},
\]

where \( y_k \) are linear approximations of given data \( D = \{(x_i, y_i)\}, x_i \in X, y_i \in Y, (i = 1, \ldots, n) \), on fuzzy intervals \( A_k \), the number of fuzzy sets \( m \) is unknown and \( t = 2 \).

Fuzzy intervals in fuzzy partitions are defined as parametric generalized bell membership functions (Jang...
Fig. 4. Example of linguistic description of dependencies in data: (a) initial fuzzy partition; (b)–(d) partitions sequentially constructed by genetic algorithm. Linguistic description of final result obtained by re-translation procedure is following: IF X IS LESS THAN 3 THEN Y IS SLOWLY INCREASING; IF X IS GREATER THAN 3 THEN Y IS SLOWLY DECREASING.
where \( c_k \) is a center of membership function, \( a_k \) is its width on the level 0.5 and \( b_k \) defines its steepness \((a_k,b_k > 0)\). Moreover, fuzzy partitions \( \{A_k\}, (k = 1,\ldots,m), (m > 1) \) of \( X = [x_l, x_r] \) have been designed such that \( c_1 = x_l, c_m = x_r \), and \( c_{k+1} = c_k + a_k + a_{k+1} \) for all \( k = 1,\ldots,m-1, (m > 2) \), where \( a_1 + 2(a_2 + \cdots + a_{m-1}) + a_m = (x_r - x_l) \). From the construction it follows that all fuzzy intervals intersected at the level 0.5.

The points of intersection of fuzzy sets in the partition are called knots. Parameters \( a_k \) and \( c_k \) are determined by knots \( x_{00} = x_l, x_{0m} = x_r, x_{0k-1}, x_{0k} \in X = [x_l, x_r], (k = 1,\ldots,m) \) as follows: \( a_1 = x_{01} - x_{00}; a_k = 0.5(x_{0k} - x_{0k-1}) \), for \( k = 2,\ldots,m-1, (m > 2) \), \( a_m = x_{0m} - x_{0m-1}, c_1 = x_l, c_k = 0.5(x_{0k} + x_{0k-1}) \), for all \( k = 2,\ldots,m-1, (m > 2) \), \( c_m = x_r \). For the definition of fuzzy partition of \( X = [x_l, x_r] \) on \( m \) intervals it is sufficient to choose \( m - 1 \) knots \( x_{01}, \ldots, x_{0m-1} \) within \( [x_l, x_r] \) and define \( m \) parameters \( b_k \) of membership functions. These parameters have been found by genetic algorithm as follows.

The vector \( s = (x_{01},\ldots,x_{0m-1}, b_1,\ldots,b_m) \) with \( 2m - 1 \), \((m > 1) \) elements is called a string. The values \( x_{00} = x_l, x_{0m} = x_r \) are fixed and not included in the string. To define a fuzzy partitions \( \{A_k\}, (k = 1,\ldots,m) \) of \( X \) on fuzzy intervals the string should satisfy the conditions: \( x_{0k-1} < x_{0k}, x_{0k} \in X = [x_l, x_r], \) and \( b_k > 0, (k = 1,\ldots,m) \). For each fuzzy interval \( A_k \) from given partition the optimal approximations \( y_k = p_kx + q_k \) of Eq. (10) have been calculated. As a result, the value of fitness function (11) for the obtained fuzzy partition and approximation functions was determined.

A goal of algorithm is to find the string \( s \) minimizing the fitness function (11). The set of strings that considered on each step of genetic algorithm is called a population. Initially a sufficiently large number of intervals \( m \) in fuzzy partitions is chosen. The genetic algorithm starts with some initial string defining the partition of \( X \) and the value of fitness function for this string is calculated. Then, a population with \( n_1 \) strings is randomly generated and the values of fitness function for each string are calculated. \( n_2 \) strings with minimal values of fitness function are selected from \( n_1 + 1 \) strings. These ‘best’ strings are called elite. The strings from the elite are used for generation of new population by means of crossover and mutation operations (Batyshrin & Wagenknecht, 2002; Goldberg, 1989).

The genetic algorithm produces a given number of generations of new populations and selections of new elites. The best string from final elite with minimal value of fitness function is considered as a solution of optimization problem for given number of fuzzy intervals in fuzzy partitions.

If the obtained optimal partition does not admissible then the merging procedure is applied and the number of fuzzy intervals in partition is decreased on the unity. The genetic algorithm and merging procedure are applied repeatedly for decreasing number of intervals in partition until the obtained optimal fuzzy partition became admissible.

The example of such sequentially obtained fuzzy partitions and correspondent approximations of given data are shown in Fig. 4.

The retranslating procedure for linguistic description of fuzzy intervals in obtained partition is similar to the procedure described in Section 2 with parameter \( a = 0.5 \). Numerical values of slopes from optimal partitions are replacing by suitable linguistic terms from Table 1.

4. Conclusions

The development of a novel type of rules was motivated by papers of Zadeh (1997, 1999) where the general scheme of computing with words (CW) is described. In CW the problem, an initial data set (IDS) and a terminal data set (TDS) are formulated linguistically. A solution of the problem is obtained as result of propagation of fuzzy constraints from IDS to TDS. The method of solution of granular differential equation discussed above may be considered as an example of such computing with words. Linguistic formulation and solution of problems give possibility of reasoning about systems on general level, which is insensitive to specific context, unnecessary details, possible errors and uncertainty in data. The set of rules may represent a knowledge base invariant to the change of parameters of some problem area or may describe the parametric family of fuzzy functions when the value of some variable is considered as (perhaps hidden) parameter.

Fuzzy logic and soft computing methods have diverse applications in petroleum industry (Fino et al., 2001; Gedeon, Tamhane, Lin, & Wong, 2001; Hu, Chan, & Huang, 2003; Nikravesh & Aminzadeh, 2001; Tamhane, Wong, & Aminzadeh, 2002). Below we consider some possible applications in this area of models discussed in the previous sections.

In Mallon and Swarbrick (2002) it is studied the variability of porosity-depth trends. The main conclusions obtained on the base of large volume of experimental data and deep theoretical analyses are described linguistically which may be used as a basis for construction of knowledge based system in considered area. This knowledge base may contain the statements like ‘The change in the rate of porosity decline appears to occur at approximately 1500 m’, ‘Rapid porosity loss’, ‘Slow porosity loss’, ‘A greater degree of porosity loss’, ‘Dominantly chemical processes such as dissolution and reprecipitation are responsible for the rapid porosity loss in the first section of the compaction trend’, ‘Increasing temperature increases the rate of chemical reactions such as dissolution and ionic diffusion
within carbonates’, etc. Such knowledge base may be partially formalized by means of granular functions described in Section 2. Such general linguistic descriptions may be used for draft evaluation of porosity value on different depth for regions with the similar geological characteristics.

The expert system with linguistic information about dependencies between input and output parameters (totally more than 30 parameters) of chemical processes was used in Batyrshin, Zakuanov, and Bikushev (1994) for qualitative modeling a chemical reactor used in petrochemical industry. The expert information like ‘An increasing temperature leads to an increasing density’ was provided by qualitative weights characterizing the intensity of influence of input parameter on output parameter. These qualitative weights were used for ordering the input parameters whose change could influence on the process in desirable direction. This expert system together with simulator of chemical reactor was used as an interactive intelligent training system for better understanding of chemical processes in petrochemical industry. The interpretation of the weights of rules as slope values like in Section 2 may be used for construction of intelligent simulator which not only gives advices about what input parameter should be changed but also evaluates the possible change of input parameter which will give the desirable change in output parameter. The method of solution of equation $Y(\chi) = B$ considered in Section 2 may be used for qualitative solution of this problem.

The concept of granular differential equations may be extended on systems of differential equations, on higher order differential equations and on partial differential equations. It would be interesting to apply this approach to oil deposit modeling where partial differential equations are traditionally used. The granular differential equations may be used for modeling reservoirs with uncertain information about rock properties.

The proposed approach to representation and processing of linguistic derivatives may be used for extension of methods of qualitative reasoning about processes and systems based on the use of the signs of derivatives (De Kleer & Brawns, 1984; Forbus, 1984; Kuipers, 1984). The use of granular derivatives will give the possibility not only to detect the change of the sign of derivatives but also to take into account the qualitative values of derivatives. It may be used for qualitative reasoning about the families of parametric processes and detecting the change of parameters which will cause a qualitative change of processes.

The models discussed in Section 3 may be applied to the modeling of dependencies between petrophysical rock parameters such as porosity—density, permeability—porosity, etc. Such dependencies were described in Finol et al. (2001) by TSK fuzzy models obtained as result of two-step procedure. On the first step the optimal clusterization of data has been received by fuzzy clustering. On the second step the parameters of approximating lines passing through the centers of clusters has been obtained by least-square approximation method. In the method discussed in Section 3 these two steps are joined together and the optimal clusterization of data is obtained as the result of the minimization of the approximation error. We suppose that our approach can give the better approximation of data and in future work we plan to apply this approach in fuzzy modeling of dependencies between petrophysical rock parameters.

Linguistic description of numerical data gives possibility to integrate knowledge-based and data-driven methods of modeling. Linguistic description of data may be used for data mining and compression of data, for detecting the changes in trends of time series, for systems monitoring, for prediction of possible failure, etc. (Gedeon et al., 2001; Kivikunnas, 1999; Loncaric, 1998; Tamhane et al., 2002). The rules containing in consequent parts the specific features of the shape of output variable may be used for granular shape analysis and recognition of patterns in data described by sequences of linguistic rules. For example, three sequentially connected rules with right sides (QUICKLY INCREASING, CONSTANT, QUICKLY DECREASING) may describe a pattern of jump of some economical or physical parameter. Such linguistic patterns may be important for decision-making system based on expert knowledge. A linguistic description of large arrays of data may be based on distributed computing and on integration of local linguistic descriptions of spatial, temporal or multivariable data. Such integration may be based on unification of fuzzy intervals and linguistic scales, on correction of local scales and descriptions, on union of local descriptions in global linguistic description, etc. Generally, linguistic rules may be considered as primitives for description of patterns by means of some grammar.

**Acknowledgements**

Research supported in parts by IMP project CDL.00006, RFBR Grant 02-01-0092 and BISC Program.

**References**


