Neurocomputing 78 (2012) 48–54

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Extracting linguistic rules from data sets using fuzzy logic and genetic algorithms

Dan Meng [a,*], Zheng Pei [b]

[a] School of Economics Information Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China
[b] School of Mathematics & Computer Engineering, Xihua University, Chengdu 610039, China

ARTICLE INFO

Available online 25 August 2011

Keywords:
Computing with Words
Linguistic rules
Fuzzy logic
Genetic algorithms

ABSTRACT

Linguistic rules in natural language are useful and consistent with human way of thinking. They are very important in multi-criteria decision making due to their interpretability. In this paper, our discussions concentrate on extracting linguistic rules from data sets. In the end, we firstly analyze how to extract complex linguistic data summaries based on fuzzy logic. Then, we formalize linguistic rules based on complex linguistic data summaries, in which, the degree of confidence of linguistic rules from a data set can be explained by linguistic quantifiers and its linguistic truth from the fuzzy logical point of view. In order to obtain a linguistic rule with a higher degree of linguistic truth, a genetic algorithm is used to optimize the number and parameters of membership functions of linguistic values. Computational results show that the proposed method is an alternative method for extracting linguistic rules with linguistic truth from data sets.

1. Introduction

An abundance of data in database is often beyond human cognition and comprehension. In real life, information is commonly transmitted through statements in natural language, which is also called as linguistic information, e.g. "about half of employees are young" seems useful and consistent with human way of thinking. Linguistic information often involves uncertainty, formally, the most appropriate realistic models for dealing with linguistic information is Computing with Words (CWW) proposed by Zadeh in [48,49]. In uncertain information processing, extracting fuzzy rules and modeling with fuzzy rule-based systems is an important aspect and has been widely researched in [5,9,23,26,29,32,35–37,40,41]. Based on fuzzy logic [45,46], modeling with fuzzy rule-based systems can be performed depending on the desired degree of interpretability and accuracy of the final model. Unfortunately, interpretability and accuracy are contradictive properties directly depending on the learning process and model structure. When modeling some complex systems, fuzzy rule-based systems process accuracy but lack interpretability in fuzzy rules described by fuzzy sets, in which, genetic algorithms and/or neural network are main tools for optimizing the number of linguistic terms, membership function parameters and/or the number of rules [2,6,8,10,14,16,17,19,20,24,25,38–40,43]. For example, by using neural network or genetic algorithms, we extract the following fuzzy rule $\tilde{R}$: *If X is* μ_A, *then Y is* μ_B. However, we do not know which linguistic terms can be used to interpret μ_A and μ_B. Differently, a linguistic rule is expressed by $\tilde{R}_l$: *If X is big*, *then Y is small*, it owns interpretability. Such linguistic rules are very important in multi-criteria decision making, new product development, *etc.*

Linguistic rule-based systems composed of linguistic variables [47] taking values in a term set with a real-world meaning possess interpretability but lack accuracy. In recent years, many different possibilities to improve the accuracy of linguistic fuzzy rule-based systems while preserving its intrinsic interpretability have been considered, e.g. Alcalá et al. propose a new postprocessing approach to perform an evolutionary lateral tuning of membership functions and obtain linguistic models with higher levels of accuracy while maintaining good interpretability in [1]. In addition, based on 2-tuples linguistic representation model, Alcalá et al. present a multi-objective evolutionary approach to quickly learn the associated rule base and generate a set of linguistic fuzzy-rule based systems with different tradeoffs between accuracy and interpretability in regression problems in [3,4]. Cordón et al. use genetic process to learn the number of linguistic terms per variable, the membership function parameters that define their semantics and the number of rules and their composition in [11]. Ishibuchi et al. provide a three-objective genetics-based machine to extract linguistic rules for high-dimensional pattern classification problems in [18]. Broekhoven

* Corresponding author.
E-mail addresses: danmeng2006@gmail.com, mengd_t@swufe.edu.cn (D. Meng), zhengpei@mail.xhu.edu.cn (Z. Pei).

0925-2312/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2011.05.029

et al. use a classic genetic algorithm with binary chromosomes, as well as a real-coded genetic algorithm to optimize the membership functions of the input variables while preserve their interpretability in fuzzy ordered classifiers in [7]. Fernandez et al. use the pairwise learning approach and preference relations to deal with multi-class classification for linguistic rule based classification systems, the method improves the performance of the linguistic rule based classification systems in [15]. Evsukoff et al. use spectral analysis with structure and parameters optimization to handle the interpretability of the rules and the model's accuracy such that it can be used as tool for data understanding in [13]. In [30], we have discussed extracting linguistic data summaries from personnel database. Linguistic data summaries is a linguistic statement investigated in [21,22,27,28,31,33,34,44], from the fuzzy logical point of view, we have analyzed membership functions of fuzzy quantifiers and linguistic truth, and provided two methods to extract simple and complex linguistic data summaries. One is based on *max* operator and the other is based on aggregation operator. To obtain a complex linguistic data summary with a higher degree of truth, we have also used genetic algorithms for optimizing the number and membership functions of linguistic terms.

Formally, linguistic rule $\tilde{R}_l$: *If X is big*, *then Y is small* is a fuzzy statement. In fuzzy logic system, every fuzzy statement is given a linguistic truth [34], e.g. very true, rather true, almost false, quite false, *etc*. In uncertain inference, the more true of fuzzy statements, the more confident of their conclusion. From the inference point of view, truth of linguistic rule can be also used to explain accuracy of linguistic rule, hence, obtaining linguistic rule with higher linguistic truth from database is desired. Obviously, truth of linguistic rule $\tilde{R}_l$: *If X is big*, *then Y is small* is determined completely once the truth of linguistic data summaries '*X is big*' and '*Y is small*' in fuzzy logic system. Hence, the following three steps are used to extract linguistic rule with linguistic truth from database:

(1) Extract (complex) linguistic data summaries with linguistic truth from database.
(2) Obtain linguistic rules based on complex linguistic data summaries.
(3) Obtain truth of linguistic rule based on truth of linguistic data summaries in fuzzy logic system.

In this paper, we provide an alternative method to extract linguistic rules with linguistic truth from decision tables based on linguistic data summaries, in which, linguistic quantifiers and linguistic truth are obtained from the fuzzy logical point of view. Genetic algorithms will be used for optimizing the number and membership functions of linguistic terms. The rest of this paper is arranged as follows: In Section 2, we make a review of linguistic data summaries. In Section 3, we formalize linguistic rules based on complex linguistic data summaries and present a method for obtaining linguistic quantifiers and linguistic truth of linguistic rules. In Section 4, we provide the objective function for optimizing the number and parameters of linguistic rules with higher fuzzy linguistic quantifier and linguistic truth based on GAs. In Section 5, we give computational results for evaluation of red wine. We conclude in Section 6.

2. Linguistic data summary

A simple linguistic data summary can be expressed, e.g. *'most of employees are young' is true*. It can be formalized by *'Qys are S' is T*, in which, Q is a fuzzy linguistic quantifier, $Y=\{y_i | i=1,\ldots,n\}$ is a set of objects, S is a summarizer (a fuzzy linguistic value) of a (an) quality (attribute) for Y, e.g. *young* is summarizer of ages of employees, and T is linguistic truth for the fuzzy statement *'Qys are S'*. Denote $D=\{v(y_i) | i=1,\ldots,n\}$ the values of quality v for objects Y, then a summarizer S of v is semantically represented by a fuzzy set $\mu_S : D \longrightarrow [0,1]$. From the logical point of view, the fuzzy sets of a fuzzy linguistic quantifier and a linguistic truth are different from the fuzzy set of a summarizer in a linguistic data summary. In fact, for the classical universal quantifier $\forall$, numbers of objects are emphasized, *i.e.*, $(\forall u)p(u)$ means "every u satisfies $p(u)$." Let $P(Y)=\{A | A\subseteq Y\}$ be the power set of Y. Define a binary relation on $P(Y)$: $A\sim B \Longleftrightarrow |A|=|B|$, where $|A|$ is the cardinality of A and "$\sim$" is an equivalence relation on $P(Y)$, denote $\overline{P}(Y)=P(Y)/\sim$. Then the fuzzy sets of Q and T can be defined as $\mu_Q : \overline{P}(Y)\longrightarrow[0,1]$ and $\mu_T : \mu_Q(\overline{P}(Y))\longrightarrow[0,1]$, respectively. Accordingly, a simple linguistic data summary can be extracted automatically at level θ as follows [30]:

- Fixing a linguistic value S (it can be one or several) and a level (threshold) θ decided by experts or users. Let

$$D_S^\theta = \mu_S^{-1}(v(y_i)) = \{v(y_i) | \mu_S(v(y_i)) \geq \theta\}. \tag{1}$$

- Selecting a fuzzy linguistic quantifier Q, *i.e.*, can be selected such that

$$\mu_Q(C) = max\{\mu_{Q_1}(C), \mu_{Q_2}(C), \ldots, \mu_{Q_m}(C)\}, \tag{2}$$

in which $C=\{y_i | v(y_i)\in D_S^\theta\}$.

- Selecting linguistic truth T, *i.e.*,

$$\mu_T(\mu_Q(C)) = max\{\mu_{T_1}(\mu_Q(C)), \mu_{T_2}(\mu_Q(C)), \ldots, \mu_{T_k}(\mu_Q(C))\}. \tag{3}$$

The so-called complex linguistic data summary has the form: *'Qys are S_1 and (or) $\cdots$ and (or) S_r' is T*, in which, S_1 is a summarizer of v_1 for Y, ..., S_r is a summarizer of v_r for Y, respectively. Based on (1), (2) and (3), we can extract simple linguistic data summaries *'Q_1ys are S_1' is T_1*, $\cdots$ and *'$Q_r$ys are S_r' is T_r*, respectively. Intuitively, extracting a complex linguistic data summary is equal to combining $\{Q_1,\ldots,Q_r\}$ and $\{T_1,\ldots,T_r\}$ to obtain Q and T, respectively.

Example 1 (*Pei et al. [30]*). Given a database (Table 1). Let $S_{\mathbf{age}}$={young (y), middle age (ma)}, $S_{\mathbf{salary}}$={low (l), high (h)}, Q={several (s), about half (ah), most (m)}, T={approximately true (at), true (t), very true (vt)}. Membership functions are given as follows:

$$\mu_y(x)=\begin{cases}1, & \text{if } x\in[25,30],\\ 4-\dfrac{x}{10}, & \text{if } x\in(30,40],\\ 0, & \text{if } x>40,\end{cases} \qquad \mu_{ma}(x)=\begin{cases}1, & \text{if } x\geq 45,\\ \dfrac{x}{10}-3.5, & \text{if } x\in(35,45),\\ 0, & \text{if } x\leq 35,\end{cases}$$

Table 1
Personnel database.

V\Y	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}
Age	25	48	31	35	28	51	37	43	34	27	53	45
Salary	1.8	2.0	2.8	3.0	2.8	3.0	2.3	2.5	3.5	2.9	3.0	3.1

$$\mu_l(x)=\begin{cases}1, & \text{if } x\in[1.8,2],\\ \frac{2.5-x}{2}, & \text{if } x\in(2,2.5],\\ 0, & \text{if } x\geq 2.5,\end{cases}\quad \mu_h(A)=\begin{cases}1, & \text{if } x\in[3.3,3.5],\\ x-2.3, & \text{if } x\in[2.3,3.3),\\ 0, & x<2.3,\end{cases}$$

$$\mu_s(A)=\begin{cases}\frac{|A|-1}{2}, & \text{if } |A|\in[1,3],\\ 2-\frac{|A|}{3}, & \text{if } |A|\in(3,6],\\ 0, & \text{if } |A|\in(6,12],\end{cases}\quad \mu_{ah}(A)=\begin{cases}0, & \text{if } |A|\in[1,4]\\ \frac{|A|}{2}-2, & \text{if } |A|\in[4,6],\\ 4-\frac{|A|}{2}, & \text{if } |A|\in(6,8],\\ 0, & \text{if } |A|\in(8,12],\end{cases}$$

$$\mu_m(A)=\begin{cases}\frac{|A|}{6}-1, & \text{if } |A|\in[6,12],\\ 0, & \text{if } |A|\in[0,6],\end{cases}\quad \mu_{at}(x)=\begin{cases}1-1.25x, & \text{if } x\in[0,0.8],\\ 0, & \text{if } x\in(0.8,1],\end{cases}$$

$$\mu_t(x)=\begin{cases}5(1-x), & \text{if } x\in[0.8,1],\\ \frac{10(x-0.5)}{3}, & \text{if } x\in[0.5,0.8),\\ 0, & \text{if } x\in[0,0.5),\end{cases}\quad \mu_{vt}(x)=\begin{cases}5x-4, & \text{if } x\in[0.8,1],\\ 0, & \text{if } x\in[0,0.8),\end{cases}$$

(1) Fixing linguistic values $s'=young\in S_{\mathbf{age}}$ and $s''=high\in S_{\mathbf{salary}}$. Let threshold $\theta=0.5$, then $D^{0.5}_{s'}=\{V(y_i)|\mu_{s'}(V(y_i))\geq 0.5\}=\{25,31,35,28,34,27\}$, $D^{0.5}_{s''}=\{V(y_i)|\mu_{s''}(V(y_i))\geq 0.5\}=\{2.8,3.0,3.5,2.9,3.1\}$, $A_{s'}=\{y_i|V(y_i)\in D^{0.5}_{s'}\}=\{y_1,y_3,y_4,y_5,y_9,y_{10}\}$ and $A_{s''}=\{y_i|V(y_i)\in D^{0.5}_{s''}\}=\{y_3,y_4,y_5,y_6,y_9,y_{10},y_{11},y_{12}\}$.

(2) According to μ_s, μ_{ah}, μ_m and $A_{s'}$, obtain $\mu_s(A_{s'})=0,\mu_{ah}(A_{s'})=1$, and $\mu_m(A_{s'})=0$, *i.e.*, $max\{\mu_s(A_{s'}),\mu_{ah}(A_{s'}),\mu_m(A_{s'})\}=\mu_{ah}(A_{s'})$, and $\mu_{at}(\mu_{ah}(A_{s'}))=\mu_t(\mu_{ah}(A_{s'}))=0,\mu_{vt}(\mu_{ah}(A_{s'}))=1$. The simple linguistic data summary is "'about half of employees are young' is very true". Similarly, we also have a simple linguistic data summary "'most of employees have high salary' is approximately true".

(3) Obtaining complex linguistic data summary "'Q employees are young and have high salary' is T", there exists the following direct method: let $C=A_{s'}\cap A_{s''}=\{y_3,y_4,y_5,y_9,y_{10}\}$, so, $\mu_s(C)=\frac{1}{3}$, $\mu_{ah}(C)=0.5$, $\mu_m(C)=0$, $\mu_{at}(\mu_{ah}(C))=1$, $\mu_t(\mu_{ah}(C))=\mu_{vt}(\mu_{ah}(C))=0$, and the complex linguistic data summary is "'about half of employees are young and have high salary' is approximately true".

3. Linguistic rules from data sets

A decision table is formalized as $(O,A\cup B,f)$, in which, O is the set of objects, $A\cup B$ is the set of attributes such that $A\cap B=\emptyset$ and A is called as the set of conditional attributes, B is called as the set of decision attributes, f is a mapping and expresses some relation between all objects and their attribute values, e.g. assume $A\cup B=\{a_1,\ldots,a_n,b_1,\ldots,b_m\}$, for any $c\in A\cup B$, denotes V_c the set of attribute values of c for all objects O, then for every $o\in O$, $f(o)=(f_{a_1}(o),\ldots,f_{a_n}(o),f_{b_1}(o),\ldots f_{b_m}(o))\in V_{a_1}\times\cdots\times V_{a_n}\times V_{b_1}\times\cdots\times V_{b_m}$. In practice, objects are also interpreted as cases, states, processes, or observations, *etc*. Attributes are interpreted as features, variables, characteristic, or conditions, *etc*. Data sets (or information systems) also called data tables, attribute-value systems, knowledge representation systems, *etc*. They are widely used for representing knowledge in artificial intelligence. In a decision table, decision rules can be extracted and formalized as "If τ, then ϕ", *i.e.*, $\tau\to\phi$, in which, τ is a formula generated by some $v_i\in\bigcup_{a\in A}V_a$ finitely using connectives $\wedge$ or $\vee$, ϕ is a formula generated by some $v_j\in\bigcup_{b\in B}V_b$ finitely using connectives $\wedge$ or $\vee$ [32], as a special case, if ϕ is a class label, then $\tau\to\phi$ is a fuzzy rule based classification systems.

Here, we consider linguistic rules from data sets, *i.e.*, in a decision rule $\tau\to\phi$, τ is a complex linguistic data summary on (O,A,f) and ϕ is a complex linguistic data summary on (O,B,f). Evidently, such linguistic rules from data sets emphasize firstly interpretability of decision rule, e.g. in a universe of discourse interval [0, 20], "if x is big, then y is small" is more understandable and interpretable in natural language than "if x is included in interval [10, 20], then y is included in interval $[0,8]$", especially, when there are imprecise data, missing values and multiple descriptors included in data sets, it needs such linguistic rules to represent decision knowledge in natural language. On the other hand, the advantage of such linguistic rules from data sets is that it does not lack the degree of confidence. In fact, in a linguistic data summary, fuzzy linguistic quantifier Q and linguistic truth T together express the degree of confidence of the linguistic data summary, intuitively, the bigger fuzzy linguistic quantifier Q and linguistic truth T are, the higher the degree of confidence of the linguistic data summary is. Hence, the degree of confidence of linguistic rule $\tau\to\phi$ from a decision table can be explained by combining degrees of confidence of τ and ϕ. Based on the above mentioned discussions, linguistic rules from data sets are formally defined as follows in this paper.

Definition 2. Let $(O,A\cup B,f)$ be a decision table. A linguistic rule from data sets is formed as $\tau\to\phi$ with a linguistic truth T, in which, τ is a complex linguistic data summary on (O,A,f) with a linguistic truth T_τ, ϕ is a complex linguistic data summary on (O,B,f) with a linguistic truth T_ϕ and T is decided by combining T_τ and T_ϕ.

To extract a linguistic rule from data sets according to complex linguistic data summaries on (O,A,f) and (O,B,f)), the following method is provided in this paper. Similar to Example 1, suppose that membership functions of fuzzy linguistic quantifier and linguistic truth are given by $\mu_{Q_1},\ldots,\mu_{Q_l}$ and $\mu_{T_1},\ldots,\mu_{T_s}$, respectively.

1. Let simple linguistic data summaries on (O,A,f) (or (O,B,f)) at level θ be "'Q_1*objects are* S_1' *is* T_1", $\cdots$ and "'Q_r*objects are* S_r' *is* T_r". $D^\theta_{S_{k_1}}=\mu^{-1}_{S_{k_1}}(f_a(o_i))=\{f_a(o_i)|a\in A,o_i\in O,\ \mu_{S_{k_1}}(f_a(o_i))\geq\theta\}$ $(1\leq k_1\leq r)$. $C_1=\bigcap^r_{k_1=1}\{o_i|f_a(o_i)\in D^\theta_{S_{k_1}}\}$, then a complex linguistic data summary on (O,A,f) (or (O,B,f)) at level θ is "'Q_τ *objects are* S_1 *and* $\cdots$ *and* S_r' *is* T_τ", in which, Q_τ and T_τ satisfy
$$\mu_{Q_\tau}(C_1)=max\{\mu_{Q_1}(C_1),\mu_{Q_2}(C_1),\ldots,\mu_{Q_l}(C_1)\},\tag{4}$$
$$\mu_{T_\tau}(\mu_{Q_\tau}(C_1))=max\{\mu_{T_1}(\mu_{Q_\tau}(C_1)),\ldots,\mu_{T_s}(\mu_{Q_\tau}(C_1))\}.\tag{5}$$
2. A linguistic rule from data sets at level θ is extracted as "'If Q_τ*objects are* S_1 *and* $\cdots$ *and* S_r, then Q_ϕ*objects are* S'_1 *and* $\cdots$ *and* S'_p' is $T_{s'}$".
3. Linguistic truth T of the linguistic rule from decision tables can be obtained by the following method:
$$\mu_{T_{s'}}(\mu_{Q_{l'}}(C))=max\{\mu_{T_1}(\mu_{Q_{l'}}(C_1)),\ldots,\mu_{T_s}(\mu_{Q_{l'}}(C))\},\tag{6}$$
$$\mu_{Q_{l'}}(C)=max\{\mu_{Q_1}(C),\mu_{Q_2}(C),\ldots,\mu_{Q_l}(C)\},\tag{7}$$
in which, $C=(\bigcap^r_{k_1=1}\{o_i|f_a(o_i)\in D^\theta_{S_{k_1}}\})\cap(\bigcap^p_{k_2=1}\{o_i|f_b(o_i)\in D^\theta_{S'_{k_2}}\})$ and
$D^\theta_{S'_{k_2}}=\mu^{-1}_{S'_{k_2}}(f_b(o_i))=\{f_b(o_i)|b\in B,o_i\in O,\mu_{S'_{k_2}}(f_b(o_i))\geq\theta\}(1\leq k_2\leq p)$.

Example 3. The decision table shown in Table 2 is used to predict human wine taste preferences. The problem has been investigated in [12], in which, there are two data sets: red wine (1599 samples) and white wine (4898 samples), 11 conditional

Table 2
Data set of red wine.

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	o_9	o_{10}	o_{11}
pH	3.51	3.39	3.52	3.38	3.37	3.4	3.44	3.16	2.93	3.42	3.23
Alcohol	9.4	10	9.7	9.8	9	9.4	10.7	9.1	9.9	10.5	9.7
Quality	5	7	5	6	4	6	5	4	6	7	7

attributes based on physicochemical tests (e.g. *pH* values, *etc.*) and 1 decision attribute based on sensory data (quality, score between 0 and 10 made by wine experts). In this example, we only select red wine data, in which, 11 samples, 2 conditional attributes (*pH* and alcohol) and 1 decision attribute (quality) are included. Define low (*L*), middle (*M*) and high (*H*) for *pH* and alcohol, bad (*B*), middle (*M*) and excellent (*E*) for quality, fuzzy linguistic quantifier and linguistic truth are similar to Example 1, e.g.

$$\mu_M^{pH}(x)=\begin{cases}\dfrac{20x}{3}-21, & x\in[3.15,3.3],\\ 5(3.5-x), & x\in(3.3,3.5],\end{cases}\quad \mu_E(x)=\begin{cases}2x-11, & x\in[5.5,6],\\ 1, & x\in(6,7],\end{cases}$$

$$\mu_M^{alcohol}(x)=\begin{cases}2x-19, & x\in[9.5,10],\\ 21-2x, & x\in(10,10.5],\end{cases}$$

at level 0.4, we have $D_M^{0.4,pH}=\{3.39,3.38,3.37,3.4,3.42,3.23\}$, $D_M^{0.4,alcohol}=\{10,9.7,9.8,9.9\}$, $D_E^{0.4}=\{6,7\}$. $A_M^{pH}=\{o_2,o_4,o_5,o_6,o_{10},o_{11}\}$, $A_M^{alcohol}=\{o_2,o_3,o_4,o_9,o_{11}\}$, $A_E=\{o_2,o_4,o_6,o_9,o_{10},o_{11}\}$. Accordingly, we can obtain the following linguistic rule from Table 2.

'If (Several) red wine has middle PH and middle alcohol, then (about half) red wine is excellent' is very true.

4. Optimization of linguistic rules based on genetic algorithms

In the above mentioned linguistic rules, the fuzzy linguistic quantifier and linguistic truth express how many objects satisfy linguistic rules and degrees of association between conditions and conclusions, respectively. In this section, we use genetic algorithms (GAs) to optimize membership functions of linguistic rules, which makes linguistic rules to inherit higher fuzzy linguistic quantifier and linguistic truth from data base.

Let there exist L attributes in $A\cup B$, and each domain of attribute is denoted by $D_l\subset R^+, l=1,\dots,L=|A\cup B|$ (the cardinality of $A\cup B$), then each object $y_i\in Y$ is understood as a point on space $D_1\times D_2\times\cdots\times D_L$, *i.e.*, $y_i=(d_{i1},d_{i2},\dots,d_{iL}), d_{il}\in D_l$. Formally, a linguistic rule is corresponding to a fuzzy class on $D_1\times D_2\times\cdots\times D_L$. Let each D_l of the space $D_1\times D_2\times\cdots\times D_L$ be partitioned into K_l fuzzy subsets $\{\mu_{k_l}^l\,|\,k_l=1,\dots,K_l\}$, then $D_1\times D_2\times\cdots\times D_L$ is divided into $K_1\times K_2\times\cdots\times K_L$ fuzzy subspaces, and each fuzzy subspace can be expressed by a linguistic *If–Then* rule at level θ:

R_g^θ : If $(Q_{R_g}^A)$ objects are $\mu_{k_1}^1$ and $\cdots$ and $\mu_{k_{|A|}}^{|A|}$, then $(Q_{R_g}^B)$ objects are

$\mu_{k_{|A|+1}}^{|A|+1}$ and $\cdots$ and $\mu_{k_L}^L$ is $T_{R_g^\theta}$.

In which, $\mu_{k_l}^l$ $(l=1,\dots,L)$ is a fuzzy subset of D_l, the fuzzy linguistic quantifiers $Q_{R_g}^A$ and $Q_{R_g}^B$ are decided by (4), linguistic truth $T_{R_g^\theta}$ is decided by (6) and (7).

The main steps for optimizing the number and parameters of membership functions using a *GAs* can be described as follows [30]:

(1) *Encoding the solution*: The two components of the solution to be encoded are the number of linguistic terms and the membership functions of linguistic terms.

1. Number of labels (S_1). In this paper, there are L variables (qualities), the number of labels per variable is stored into an integer array of length L. In this paper, the possible values considered are the set $\{3,5,7,9\}$.
2. Membership functions (S_2). In this paper, we deal with triangular functions only. A real number array of $L\times 9\times 3$ positions is used to store the membership functions. Of course, if a chromosome does not have the maximum number of labels in one variable, the space reserved for the values of these labels is ignored in the evaluation process.

If s_l is the granularity of variable l $(l=1,\dots,L)$, $s_l\in\{3,5,7,9\}$, $P_{lj}^1,P_{lj}^2,P_{lj}^3$ are the definition points of the label j of the variable l, and S_{2l} is the information about the fuzzy partition of variable l in S_2, then a graphical representation of the chromosome is shown as follows:

$$S_1=(s_1,s_2,\dots,s_L),\quad S_{2l}=(P_{l1}^1,P_{l1}^2,P_{l1}^3,\dots,P_{ls_l}^1,P_{ls_l}^2,P_{ls_l}^3),$$

$$S_2=(S_{21},S_{22},\dots,S_{2L}),\quad S=S_1S_2.$$

Uniform fuzzy partitions are denoted by $(V_{lj}^1,V_{lj}^2,V_{lj}^3)$ for each variable. Variation intervals defined for each one of membership functions are [11]

$$P_{lj}^1\in[L_{lj}^1,R_{lj}^1]=\left[V_{lj}^1-\frac{V_{lj}^2-V_{lj}^1}{2},\quad V_{lj}^1+\frac{V_{lj}^2-V_{lj}^1}{2}\right],$$

$$P_{lj}^2\in[L_{lj}^2,R_{lj}^2]=\left[V_{lj}^2-\frac{V_{lj}^2-V_{lj}^1}{2},\quad V_{lj}^2+\frac{V_{lj}^3-V_{lj}^2}{2}\right],$$

$$P_{lj}^3\in[L_{lj}^3,R_{lj}^3]=\left[V_{lj}^3-\frac{V_{lj}^3-V_{lj}^2}{2},\quad V_{lj}^3+\frac{V_{lj}^3-V_{lj}^2}{2}\right].$$

(2) *Initializing gene pool*: The initial population is composed of four groups:

1. In the first group, each chromosome will have the same number of labels in all its variables and the membership functions are uniformly distributed across the domain of variable.
2. In the second group, each chromosome can have a different granularity per variable (different values in S_1) and the membership functions are uniformly distributed as in the first part.
3. In the third group, each chromosome will have the same number of labels in all its variables. Then a uniform fuzzy partition is built for each variable as in the first group and the variation intervals of all the definition points are calculated. Finally, a value for all the definition points is randomly chosen from the correspondent variation interval.
4. In the last group, each chromosome can have different numbers of labels per variable as in second group and the membership functions are calculated in the same way as in the third group, a random value is in the variation interval.

(3) *Evaluating the chromosome*: Maximizing degrees of fuzzy linguistic quantifier and linguistic truth at level θ are the goals, so, we have

$$\mathrm{Min}: f(s) = \sum_{g=1}^{G} (w_1 \times s' + w_2 \times \mu_{T_{s'}}(\mu_{Q_{l'}}(C)) + w_3 \times l' + w_4 \times \mu_{Q_{l'}}(C)),$$

where s is a chromosome, G is the number of linguistic rules extracted at level θ for s, $\mu_{T_{s'}}(\mu_{Q_{l'}}(C))$ and $\mu_{Q_{l'}}(C)$ are decided by (6) and (7), w_1, w_2, w_3 and w_4 are weights fixed by users. The objective function $f(s)$ is treated as the fitness function in *GAs*.

(4) *Genetic operators*: Since there is a strong relationship among the two chromosome parts, operators working cooperatively in S_1 and S_2 are required in order to make best use of the representation used.

- *Selection*: Let current population Ψ. The selection probability $P(s)$ of chromosome s is

$$P(s) = (f_{max}(\Psi) - f(s)) \Big/ \sum_{s' \in \Psi} (f_{max}(\Psi) - f(s')), \tag{8}$$

in which $f_{max}(\Psi) = max\{f(s) | s \in \Psi\}$.
- *Crossover*: Two different crossover operators are considered depending on the scope of the two parents [42]:

1. Crossover when both parents have the same granularity level per variable, in this case, the crossover operator in S_2 and obviously, by maintaining the parent S_1 values in the offspring. If $(S_2^v)^t = ((P_{11}^1)^v, \ldots, (P_{LS_L}^3)^v)$ and $(S_2^w)^t = ((P_{11}^1)^w, \ldots, (P_{LS_L}^3)^w)$ are to be crossed, the follows four offspring are generated, in which, $i = 1, 2, 3$:

$$(S_2^{vw})_1^{t+1} = ((P_{11}^1)^{vw}, \ldots, (P_{LS_L}^3)^{vw}), \quad (P_{ls_l}^i)^{vw} = d(P_{ls_l}^i)^v + (1-d)(P_{ls_l}^i)^w,$$

$$(S_2^{vw})_2^{t+1} = ((P_{11}^1)^{vw}, \ldots, (P_{LS_L}^3)^{vw}), \quad (P_{ls_l}^i)^{vw} = (1-d)(P_{ls_l}^i)^v + d(P_{ls_l}^i)^w,$$

$$(S_2^{vw})_3^{t+1} = ((P_{11}^1)^{vw}, \ldots, (P_{LS_L}^3)^{vw}), \quad (P_{ls_l}^i)^{vw} = max\{(P_{ls_l}^i)^v, (P_{ls_l}^i)^w\},$$

$$(S_2^{vw})_4^{t+1} = ((P_{11}^1)^{vw}, \ldots, (P_{LS_L}^3)^{vw}), \quad (P_{ls_l}^i)^{vw} = min\{(P_{ls_l}^i)^v, (P_{ls_l}^i)^w\}.$$

This operator uses a parameter that is either a constant or a variable whose value depends on the age of the population [42]. The resulting descendants are the two best of the four possible offspring combinations.

2. Crossover when the parents encode different granularity levels. Let $S^v = ((s_1)^v, \ldots, (s_l)^v, (s_{l+1})^v, \ldots, (s_L)^v, (S_{21})^v, \ldots, (S_{2\,l})^v, (S_{2(l+1)})^v, \ldots, (S_{2L})^v)$, $S^w = ((s_1)^w, \ldots, (s_l)^w, (s_{l+1})^w, \ldots, (s_L)^w, (S_{21})^w, \ldots, (S_{2l})^w, (S_{2(l+1)})^w, \ldots, (S_{2L})^w)$ be crossed at point l, the two resulting offspring are $S_1^{vw} = ((s_1)^v, \ldots, (s_l)^v, (s_{l+1})^w, \ldots, (s_L)^w, (S_{21})^v, \ldots, (S_{2\,l})^v, (S_{2(l+1)})^w, \ldots, (S_{2L})^w)$ and $S_2^{vw} = ((s_1)^w, \ldots, (s_l)^w, (s_{l+1})^v, \ldots, (s_L)^v, (S_{21})^w, \ldots, (S_{2l})^w, (S_{2(l+1)})^v, \ldots, (S_{2L})^v)$.

- *Mutation*: Two different operators are used in this paper. (1) Mutation on S_1, in this case, once a new value $s_l' \in \{3,5,7,9\}$ at point l of S_1 is selected, a uniform fuzzy partition for this variable is stored in its corresponding zone of S_2; (2) mutation on S_2: let $(S_2^v)^t = ((P_{11}^1)^v, \ldots, (P_{ls_l}^i)^v, \ldots, (P_{LS_L}^3)^v)$ and the element $(P_{ls_l}^i)^v$ was selected for this mutation (the domain of $(P_{ls_l}^i)^v$ is $[(P_{ls_l}^i)_l^v, (P_{ls_l}^i)_r^v]$), the result is a vector $(S_2^v)^{t+1} = ((P_{11}^1)^v, \ldots, ((P_{ls_l}^i)^v)', \ldots, (P_{LS_L}^3)^v)$, and

$$((P_{ls_l}^i)^v)' = \begin{cases} (P_{ls_l}^i)^v + \Delta(t, (P_{ls_l}^i)_r^v - (P_{ls_l}^i)^v) & \text{if } e = 0, \\ (P_{ls_l}^i)^v + \Delta(t, (P_{ls_l}^i)^v - (P_{ls_l}^i)_l^v) & \text{if } e = 1. \end{cases}$$

with t being the current generation, e a random number that may have a value of zero or one, each with a probability 0.5, and $\Delta(t,y) = y(1 - r^{(1-t/T)^b})$ [11] with r being a random number drawn from a uniform distribution on the interval [0,1], T the maximum number of generations and b a parameter chosen by the user.

Table 3
The parameters of attributes.

	C	D	pH	A	Q
S_1	$\{3,5,7,9\}$	$\{3,5,7,9\}$	$\{3,5,7,9\}$	$\{3,5,7,9\}$	$\{3\}$
S_2	$(P_{1j}^1, P_{1j}^2, P_{1j}^3)$	$(P_{2j}^1, P_{2j}^2, P_{2j}^3)$	$(P_{3j}^1, P_{3j}^2, P_{3j}^3)$	$(P_{4j}^1, P_{4j}^2, P_{4j}^3)$	$(P_{5j}^1, P_{5j}^2, P_{5j}^3)$
θ	0.4				
W	$w_1 = w_3 = 0.3$, $w_2 = w_4 = 0.2$				

5. Computational results

In this section we report the results of computational experiments carried out with our proposed GAs described in the previous section. The database obtained from [12] is used to predict human red wine taste preferences. A total of 1599 samples are included in the database, 11 conditional attributes based on physicochemical tests (e.g. *pH* values, *etc.*) and 1 decision attribute based on sensory data [quality, score between 0 (very bad) and 10 (very excellent) made by wine experts to explain]. Eleven conditional attributes are fixed acidity (*FA*), volatile acidity (*VA*), citric acid (*CA*), residual sugar (*RS*), chlorides (*C*), free sulfur dioxide (*FSD*), total sulfur dioxide (*TSD*), density (*D*), pH, sulphates (*S*) and alcohol (*A*), in which, we consider that *FA*, *VA*, *CA*, *FSD*, *TSD* and pH are correlated, *C* and *S* are correlated, *RS* and *A* are correlated. Hence, we select *C*, *D*, pH and *A* as conditional attributes, quality (*Q*) as decision attribute to extract linguistic rules in our experiments. The parameters are shown in Table 3. The performance are shown in Fig. 1. In Fig. 1, the green bold line represents the final optimization results of membership function, the red hairline represents the training process of optimization. Algorithms are as follows:

Algorithm 1. The generation of membership functions.

```
Input Data, t;
  [m, n] = size(Data);
  for i = 1 : n
    for k = 1 : t(i)
        for j = 1 : m
            y(j,k,i) = y(Data)
        end
    end
  end
```

Algorithm 2. The objects with parameters more than 0.4.

```
for i = 1 : n;
    y(: , : ,i) > q;
    [m(i),n(i)] = find y(: , : ,i) > q
end
```

Algorithm 3. The optimized linguistic rules.

```
for i = 1 : n;
  [m(i),n(i)] = find y(: , : ,i) > q
  yy(i) = [m(i),n(i)]
end
  for i = 1 : t(1)
    for j = 1 : t(2)
      for p = 1 : t(3)
        for c = 1 : t(4)
          for d = 1 : t(5)
zhenghe = [yy1(find (yy1(: ,2) = = i)); yy2(find (yy2(: ,2) = = j));
            yy3(find (yy3(: ,2) = = p)); yy4(find (yy4(: ,2) = = c));
            yy5(find (yy5(: ,2) = = d))]';
```

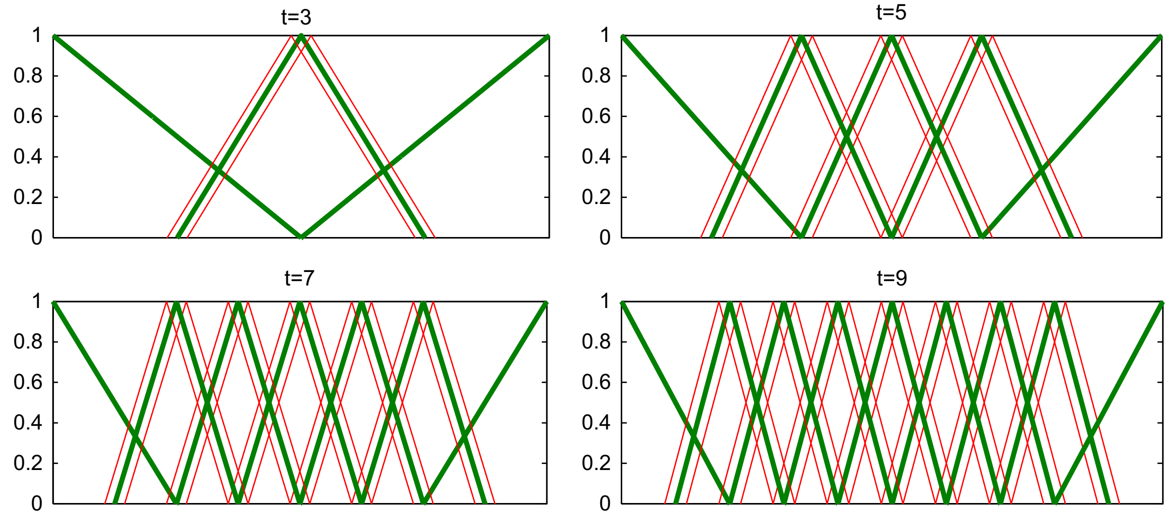

Fig. 1. The training process of the data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

```
            count = hist(zhenghe,unique(zhenghe));
            if count ~ = n
               yyy(d,c,p,j,i) = 0;
            else
               bb = find(count = = 5);
               bbb = size(bb,2);
               yyy(d,c,p,j,i) = bbb;
              end
             end
            end
           end
          end
        end
```

6. Conclusion

In this paper, based on complex linguistic data summaries, we provide a method for extracting linguistic rules from data sets, in which, the degree of confidence of linguistic rules from a data set can be explained by linguistic quantifiers and its linguistic truth from the fuzzy logical point of view. We also use genetic algorithm to optimize the number and parameters of membership functions of linguistic values, optimized linguistic rules have higher fuzzy linguistic quantifier and linguistic truth. In experiment, we use data set for evaluation of red wine to extract linguistic rules, the such linguistic rules own interpretability in natural language when they are used for evaluating the quality of red wine.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of China (60875034), the Key Laboratory of Finance Intelligence and Engineering of Sichuan Province and Economics, the grant from the project 211 (Phase III) of the Southwestern University of Finance, the research fund of Sichuan Key Laboratory of Intelligent Network Information Processing (SGXZD1002-10) and the Key Laboratory of the Radio Signals Intelligent Processing (Xihua University, XZD0818-09).

References

[1] R. Alcalá, J. Alcalá-Fdez, F. Herrera, A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection, IEEE Trans. Fuzzy Sys. 15 (4) (2007) 616–635.

[2] J. Alcalá-Fdez, R. Alcalá, M.J. Gacto, F. Herrera, Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms, Fuzzy Sets Syst. 160 (2009) 905–921.

[3] R. Alcalá, P. Ducange, F. Herrera, B. Lazzerini, F. Marcelloni, A multi-objective evolutionary approach to concurrently learn rule and data bases of linguistic fuzzy-rule-based systems, IEEE Trans. Fuzzy Syst. 17 (5) (2009) 1106–1122.

[4] R. Alcalá, J. Alcalá-Fdez, F. Herrera, J. Otero, Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation, Int. J. Approx. Reasoning 44 (2007) 45–64.

[5] N. Barakat, A.P. Bradley, Rule extraction from support vector machines: a review, Neurocomputing 74 (2010) 178–190.

[6] F.J. Berlanga, A.J. Rivera, M.J. Del Jesus, F. Herrera, GP-COACH: Genetic Programming-based learning of COmpact and ACcurate fuzzy rule-based classification systems for High-dimensional problems, Inf. Sci. 180 (2010) 1183–1200.

[7] E.V. Broekhoven, V. Adriaenssens, B.D. Baets, Interpretability-preserving genetic optimization of linguistic terms in fuzzy models for fuzzy ordered classification: an ecological case study, Int. J. Approx. Reasoning 44 (2007) 65–90.

[8] J. Casillas, O. Cordón, M.J. Del Jesus, F. Herrera, Genetic feature selection in a fuzzy rule-based classification system learning process for high-dimensional problems, Inf. Sci. 136 (2001) 135–157.

[9] T. Chavarnakul, D. Enke, A hybrid stock trading system for intelligent technical analysis-based equivolume charting, Neurocomputing 72 (2009) 3517–3528.

[10] S. Chen, C. Huang, A new approach to generate weighted fuzzy rules using genetic algorithms for estimating null values, Expert Syst. Appl. 35 (2008) 905–917.

[11] O. Cordón, F. Herrera, P. Villar, Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base, IEEE Trans. Fuzzy Syst. 9 (4) (2001) 667–674.

[12] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, J. Reis, Modeling wine preferences by data mining from physicochemical properties, Decis. Support Syst. 47 (4) (2009) 547–553.

[13] A.G. Evsukoff, S. Galichet, B.S.L.P. deLima, N.F.F. Ebecken, Design of interpretable fuzzy rule-based classifiers using spectral analysis with structure and parameters optimization, Fuzzy Sets Syst. 160 (2009) 857–881.

[14] A. Fernández, M. José del Jesus, F. Herrera, Hierarchical fuzzy rule based classification systems with genetic rule selection for imbalanced data-sets, Int. J. Approx. Reasoning 50 (2009) 561–577.

[15] A. Fernández, M. Calderón, E. Barrenechea, H. Bustince, F. Herrera, Solving multi-class problems with linguistic fuzzy rule based classification systems based on pairwise learning and preference relations, Fuzzy Sets Syst. 161 (2010) 3064–3080.

[16] H. Guo, K. Zhu, S. Gao, Y. Li, J. Zhou, Extracting fuzzy rules based on fusion of soft computing in oil exploration management, Expert Syst. Appl. 36 (2009) 2081–2087.

[17] T. Hong, C. Chen, Y. Lee, Y. Wu, Genetic-Fuzzy data mining with divide-and-conquer strategy, IEEE Trans. Evol. Comput. 12 (2) (2008) 252–265.

[18] H. Ishibuchi, T. Nakashima, T. Murata, Three-objective genetics-based machine learning for linguistic rule extraction, Inf. Sci. 136 (2001) 109–133.

[19] C. Juang, C. Lee, A fuzzified neural fuzzy inference network for handling both linguistic and numerical information simultaneously, Neurocomputing 71 (2007) 342–352.

[20] V.G. Kaburlasos, S.E. Papadakis, A granular extension of the fuzzy-ARTMAP (FAM) neural classifier based on fuzzy lattice reasoning (FLR), Neurocomputing 72 (2009) 2067–2078.
[21] J. Kacprzyk, R.R. Yager, Linguistic summaries of data using fuzzy logic, Int. J. Gen. Syst. 30 (2001) 133–154.
[22] J. Kacprzyk, S. Zadrożny, Computing with words is an implementable paradigm: fuzzy queries, linguistic data summaries, and natural-language generation, IEEE Trans. Fuzzy Syst. 18 (3) (2010) 461–472.
[23] U.M. Kaczmara, W. Trelak, Fuzzy logic and evolutionary algorithmtwo techniques in rule extraction from neural networks, Neurocomputing 63 (2005) 359–379.
[24] R.J. Kuo, S.M. Hong, Y. Lin, Y.C. Huang, Continuous genetic algorithm-based fuzzy neural network for learning fuzzy IF–THEN rules, Neurocomputing 71 (2008) 2893–2907.
[25] B. Liu, C. Chen, J. Tsao, Design of adaptive fuzzy logic controller based on linguistic-hedge concepts and genetic algorithms, IEEE Trans. Syst. Man Cybern. 31 (1) (2001) 32–53.
[26] F.D. Martino, V. Loia, S. Sessa, Fuzzy transforms method in prediction data analysis, Fuzzy Sets and Systems, ⟨10.1016/j.fss.2010.11.009⟩.
[27] A. Niewiadomski, A type-2 fuzzy approach to linguistic summarization of data, IEEE Trans. Fuzzy Syst. 16 (1) (2008) 198–212.
[28] A. Niewiadomski, On finity, countability, cardinalities, and cylindric extensions of type-2 fuzzy sets in linguistic summarization of databases, IEEE Trans. Fuzzy Syst. 18 (3) (2010) 532–545.
[29] B. Niu, Y. Zhua, X. He, H. Shen, A multi-swarm optimizer based fuzzy modeling approach for dynamic systems processing, Neurocomputing 71 (2008) 1436–1448.
[30] Z. Pei, Y. Xu, D. Ruan, K. Qin, Extracting a complex linguistic data summaries from personnel database via linguistic aggregation of simple ones, Inf. Sci. 179 (2009) 2325–2332.
[31] Z. Pei, Y. Du, L. Yi, Y. Xu, Obtaining a Complex Linguistic Data Summaries from Database based on a New Linguistic Aggregation Operator, Lecture Notes in Computer Science, vol. 3512 (IWANN 2005), Springer-Verlag, Berlin Heidelberg, 2005, pp. 771–778.
[32] Z. Pei, G. Resconi, A.J. Van Der Wal, K. Qin, Y. Xu, Interpreting and extracting fuzzy decision rules from fuzzy information systems and their inference, Inf. Sci. 176 (2006) 1869–1897.
[33] Z. Pei, D. Ruan, Y. Xu, J. Liu, Gathering linguistic information in distributed intelligent agent on the internet, Int. J. Intell. Syst. 22 (2007) 435–453.
[34] Z. Pei, D. Ruan, J. Liu, Y. Xu, Linguistic Values Based Intelligent Information Processing: Theory, Methods, and Application, Atlantis Computational Intelligence Systems, vol. 1, Atlantis Press & World Scientific, 2009.
[35] Z. Pei, X. Liu, L. Zou, Extracting association rules based on intuitionistic fuzzy sets, Int. J. Inno. Comput. Inf. Control 6 (6) (2010) 2567–2580.
[36] M. Ragulskis, K. Lukoseviciute, Non-uniformat tractor embedding for time series forecasting by fuzzy inference systems, Neurocomputing 72 (2009) 2618–2626.
[37] G. Raschia, N. Mouaddib, SAINTETIQ: a fuzzy set-based approach to database summarization, Fuzzy Sets Syst. 129 (2002) 137–162.
[38] W. Romao, A.A. Freitas, I.M.S. Gimenes, Discovering interesting knowledge from a science and technology database with a genetic algorithm, Appl. Soft Comput. 4 (2004) 121–137.
[39] L. Sanchez, I. Couso, J. Casillas, Genetic learning of fuzzy rules based on low quality data, Fuzzy Sets Syst. 160 (2009) 2524–2552.
[40] H. Song, C. Miao, Z. Shen, Y. Miao, B. Lee, A fuzzy neural network with fuzzy impact grades, Neurocomputing 72 (2009) 3098–3122.
[41] H. Song, C. Miao, Z. Shen, W. Roel, M. D'Hondt, C. Francky, A probabilistic fuzzy approach to modeling nonlinear systems, Neurocomputing 74 (2009) 1008–1025.
[42] F. Herrera, M. Lozano, J.L. Verdegay, Fuzzy connectives based crossover operators to model genetic algorithms population diversity, Fuzzy Sets Syst. 92 (1997) 21–30.
[43] X. Wang, C. Dong, T. Fan, Training T-S norm neural networks to refine weights for fuzzy if–then rules, Neurocomputing 70 (2007) 2581–2587.
[44] R.R. Yager, A new approach to the summarization of data, Inf. Sci. 28 (1982) 69–86.
[45] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[46] L.A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Syst. Man Cybern. 3 (1973) 28–44.
[47] L.A. Zadeh, The concept of linguistic variable and its application to approximate reasoning. Part 1, 2, 3, Inf. Sci. 8 (9) (1975) 199–249 301–357, 43–80.
[48] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Trans. Fuzzy Syst. 4 (2) (1996) 103–111.
[49] L.A. Zadeh, From computing with numbers to computing with words—from manipulation of measurements to manipulation of perceptions, IEEE Trans. Circ. Syst. 45 (1999) 105–119.

Dan Meng is an associate professor in School of Economics Information Engineering, Southwestern University of Finance and Economics. She received her PhD from Southwest Jiaotong University in 2004. Her research topic includes: computing with words, uncertainty reasoning, computational intelligence, etc.

Zheng Pei is a professor in School of Mathematics & Computer Engineering, Xihua University. He finished his post-doctor research in 2004. He received his PhD from Southwest Jiaotong University in 2002. His research topic includes: rough sets theory, fuzzy sets theory, logical reasoning and linguistic information processing and so on.