A Novel Distance Measure of Multi-Granularity Linguistic Variables and Its Application to MADM

Mei Cai, Zaiwu Gong, Jie Cao, and Minjie Wu

Abstract

Many decision problems are under uncertain environments with vague and imprecise information using multi-granularity linguistic variables. In this paper, we describe the linguistic hierarchical structure in a different way. The suitable numerical scales are given with the purpose of making transformation between multi-granularity linguistic variables and numerical values. A novel distance measure between multi-granularity linguistic variables is proposed. Its advantage is to solve problems of linguistic variables with different semantics. Then we develop a maximizing deviation method to determine the optimal relative weights of attributes under linguistic environment where preferences are labels in different levels of linguistic hierarchy. Application of the method is illustrated in a case study on medical diagnosis.

Keywords: Distance measure, hierarchical structure, multiple attribute decision making (MADM).

1. Introduction

In our daily lives, we continually face with natural language to describe objects and instances in order to convey the information we intend. Linguistic variables are the appropriate tools to describe vague concepts in natural language. Making decision with linguistic information is an important application. As a methodology, computing with words (CWW) provides a foundation for a computation theory of perceptions or linguistic descriptions [1], which can bring closer the gap between human’s brain mechanisms and the machine’s processes. The use of fuzzy sets is central to CWW as they provide means of modeling vagueness underlying most natural language terms [2]. Fuzzy set theory is originally proposed by Zadeh [3] and is used to represent the linguistic label semantic. So linguistic variables are often represented as labels belonging to a linguistic term set and have the relation with fuzzy subsets.

The linguistic computational approaches based on extension principle usually use ranking functions to order the fuzzy numbers and to obtain a final numerical evaluation [4, 5]. But, in some practical applications, the determination of membership functions or fuzzy sets associated with linguistic labels is difficult or impossible. So, these approaches based on fuzzy set or membership function associated with each linguistic label have their restrictions. Linguistic symbolic computational models propose another type of research methods [6-12]. Since the 2-tuple linguistic computational model [7] is proposed, many extensions to the 2-tuple linguistic model are developed. Wang and Hao [13] assign canonical characteristic values of the corresponding linguistic labels in the process of aggregating linguistic information. Xu [9] extends a discrete term set to a continuous term set by virtual linguistic terms which can be obtained and manipulated to avoid loss of information. Fan, Yue, Feng and Liu [12] model the linguistic uncertainty directly by a fuzzy relation on the set of linguistic labels. Wei [14] applies the 2-tuple linguistic representation to aggregate the linguistic assessment information in decision. In addition, Yager [15], Xu [9], [16] extend the classical families of ordered weighted averaging and geometric operators for linguistic information.

Decision making can be seen as the final outcome of some mental processes that lead to the selection of an alternative among several different ones. The decision support systems based on CWW can ease the decision makers to reach a solution. Liu, Martínez, Wang, Rodriguez and Novozhilov [17] present a comprehensive overview of currently known applications of CWW in risk assessment. Martínez and Herrera [18] give a bird’s eye view about linguistic decision making, and their applications show that the 2-tuple linguistic representation...
A novel distance measure of multi-granularity linguistic variables is proposed and its application to MADM is researched in this paper. The rest of this paper is organized as follows. In Section 2, we introduce some useful concepts and definitions. In Section 3, we propose the concept of the 2-dimension numerical scale for ordinal linguistic terms term set and set suitable numerical scale with the purpose of taking account the vagueness of linguistic terms. Hierarchical structure is proposed for a multi-granularity linguistic context. And we provide a novel distance measure of multi-granularity linguistic variables for linguistic hierarchical structure. In Section 4, we construct a non-linear optimization model to derive the weight vector of attributes in decision making. Finally a numerical example is given in section 5.

2. Preliminaries

2.1 Linguistic Variables

The concept of linguistic variables is first introduced by Zadeh [30], where it is defined as a quintuple \((L, T(L), U, S, M)\), in which \(L\) is the name of the variable, \(T(L)\) is the term set of labels or words of \(L\), \(U\) is the universe of discourse, \(S\) is the syntactic rule, and \(M\) is the semantic rule which associates with each linguistic label.

We will use an ordered linguistic term set with \(s_0 < s_1 < \cdots < s_g\) to represent a linguistic variable \((L, T(L), U, S, M)\). In Fig. 1 we can see an example of the linguistic variable ‘Height’, whose corresponding linguistic term set is \(T(\text{Height}) = \{\text{None(N), Very Low(VL), Low(L), Medium(M), High(H), Very High(VH), Perfect(P)}\}\).

![Figure 1. A set of seven terms.](image_url)

2.2 Operational laws of computing with words based on symbolic models

Several methods have been proposed for dealing with linguistic information. They are the methods based on extension principle and the methods based on symbols. The first one makes operations on the fuzzy numbers that support the semantics of the linguistic terms. The second one makes computations on the indexes of the linguistic terms. Herrera and Martinez [7] represent the linguistic information by means of 2-tuples \((s_i, \alpha) \in \widetilde{S}\), where \(s_i \in S\), and \(\alpha \in [-0.5, 0.5]\). 2-tuples linguistic representation model which belongs to the symbolic models has
been widely used in decision making problems.

**Definition 1** [7]: Let \( \mathbf{S} = \{s_0, s_1, \ldots, s_g\} \) be a linguistic term set and \( \beta \in [0, g] \) a value representing the result of a symbolic aggregation operation, then the 2-tuple that expresses the equivalent information to \( \beta \) is obtained with the following function:

\[
\Delta : [0, g] \rightarrow \mathbb{R} \times [-0.5, 0.5),
\]

with \( \Delta(\beta) = (s_i, \alpha), i = \text{round}(\beta) \)

\( \alpha = \beta - i \)

**Definition 2** [7]: Let \( \mathbf{S} = \{s_0, s_1, \ldots, s_g\} \) be a linguistic term set and \( (s_i, \alpha) \) be a 2-tuple, \( \beta \in [0, g] \). There is always a \( \Delta^{-1} \) function:

\[
\Delta^{-1} : \mathbb{R} \times [-0.5, 0.5) \rightarrow [0, g]
\]

with \( \Delta^{-1}(s_i, \alpha) = i + \alpha = \beta \).

The fuzzy linguistic representation model with 2-tuples has defined the functions that transform numerical values into a 2-tuple and vice versa without losing information, therefore, any numerical aggregation operator can be easily extended for dealing with linguistic 2-tuples.

Wang and Hao [13] provide a new version of the 2-tuple fuzzy linguistic representation model which is represented by ordinal proportional 2-tuples as \( (\alpha s_i, (1-\alpha)s_{i+1}) \) in \( \mathbf{S} \). The notion of proportional 2-tuple allows experts to express their opinions not just using one ordinal, as is normally the case, but spreading that opinion using two adjacent ordinals. The functions which transform numerical values into a proportional 2-tuples and vice versa are given.

**Definition 3** [13]: Define \( \pi : \mathbb{S} \rightarrow [0, n] \) by

\[
\pi(\alpha s_i, (1-\alpha)s_{i+1}) = i + (1-\alpha)
\]

\( \pi^{-1}(x) = ((1-\beta)s_i, \beta s_{i+1}) \)

where \( i = E(x) \), \( E \) is the integral part function, \( \beta = x - i \)

Under their proportional 2-tuple fuzzy linguistic representation model contexts, these aggregation operators not only can operate in a computational stage for CWW without any loss of information in some sense, but also can deal with linguistic labels, which do not have to be symmetrically distributed around a medium label and without the traditional requirement of having “equal distance” between them.

Xu [9] proposes the concept of virtual linguistic terms where a discrete term set is extended into a continuous term set. He calls \( s_a \in \mathbb{S} = \{s_a, s_1, s_2, \ldots, s_g\} \) the original linguistic term. And he calls \( s_a \in \mathbb{S} = \{s \leq s_a \leq s, \alpha \in [0, 1]\} \) the virtual linguistic term. Consider any two linguistic terms \( s_a, s_b \in \mathbb{S} \), and \( \mu, \mu_1, \mu_2 \in [0,1] \), some operational laws are as follows:

1. \( (s_a)^\mu = s_a^\mu \)
2. \( (s_a)^\mu \odot (s_a)^\nu = (s_a)^{\mu+\nu} \)
3. \( (s_a \odot s_b)^\mu = (s_a)^\mu \odot (s_b)^\mu \)

Dong, Xu and Yu [31] define the concept of numerical scale \( NS : S \rightarrow R \) which is the numerical index of \( s_i \) and extend the 2-tuple fuzzy linguistic representation models under the numerical scale.

**Definition 4** [31]: Let \( (s_i, \alpha) \in S \), we define the numerical scale \( NS : S \rightarrow R \)

\[
NS((s_i, \alpha)) = \left\{NS(s_i) + \alpha \times (NS(s_{i+1}) - NS(s_i)) \right\}_{\alpha \geq 0}
\]

\[
NS((s_i, \alpha)) = \left\{NS(s_i) + \alpha \times (NS(s_i) - NS(s_{i-1})) \right\}_{\alpha < 0}
\]

The symbolic models have received a quite good acceptance in the specialized literature and some applications to decision making have been researched [17, 32-38]. Reasoning mechanisms that are able to map inputs words, perceptions and propositions to real numbers have attracted much attention, notably in fuzzy logic and soft computing. But these models don’t mention the methods of solving multi-granularity linguistic computational problems effectively.

### 3. Distance Measure of Multi-Granularity Linguistic Variables

#### 3.1 Distance measure of linguistic variables with the same semantic

Some traditional distance functions are used to measure the proximity between the linguistic variables, by comparing the membership degrees vectors associated to them. Suppose \( A \) and \( B \) are two linguistic variables. They can be represented by fuzzy subsets with membership functions \( \mu_A(x) \) and \( \mu_B(x) \)

**Hamming distance**

\[
d_H(A, B) = \int \left| \mu_A(x) - \mu_B(x) \right| dx
\]

**Euclidean distance**

\[
e_E(A, B) = \sqrt{\int (\mu_A(x) - \mu_B(x))^2 dx}
\]

If the membership functions are unknown, other methods based on symbols make comparison on the indexes of the linguistic terms. Following distance measure has been proposed based on symbols:

**2-tuple symbolic index distance**

\[
m_s(A, B) = |a - b|
\]

where \( A = s_a, B = s_b, s_a, s_b \in S = \{s_0, s_1, \ldots, s_g\} \).

#### 3.2 Distance measure of multi-granularity linguistic variables

It seems difficult to accept that all individuals should
agree on the same ‘granularity of uncertainty’ in a linguistic context. In this section, we deal with multi-granularity linguistic distance measure problems.

If we know the membership functions of the two linguistic variables, above measure functions (6)-(7) can still be used. But if we measure the distance based on symbolic models, there are some changes.

Firstly, a basic linguistic term set (BLTS), \( ST = \{ s_0, s_1, \ldots, s_n \} \), has to be selected. To do this it seems reasonable to impose a granularity high enough to maintain the uncertainty degrees associated to each one of the possible domains to be unified. Then multi-granularity transformation function is applied to translate linguistic variables into labels defined on possible domains.

In linguistic approach, the linguistic term set in hierarchical structure and \( R \) be a real number set. We define the 2-scale numerical function \( 2 \cdot SNF : S^g \to R \), which is constituted by two parts: order function and vagueness function \( (O,V) \). \( (O,V) \) should satisfy these conditions:

1) In order to normalize the values of labels in different level, we require \( O \in [0,1] \) and \( V \in [0,1] \);

2) If the linguistic term \( A \) is vaguer than the term \( B \), then \( V(A) > V(B) \).

When linguistic term can be described by fuzzy set or membership function. \( (O,V) \) is defined as \( (x_\mu(s_i), \sigma_\mu(s_i)) \). \( x_\mu(s_i) \) and \( \sigma_\mu(s_i) \) are presented as following by membership function \( \mu_\xi(x) \) [40]:

\[
x_\mu(s_i) = \frac{\int x \mu_\xi(x) dx}{\int \mu_\xi(x) dx}
\]

\[
\sigma_\mu(s_i) = \left[ \frac{\int x^2 \mu_\xi(x) dx}{\int \mu_\xi(x) dx} - [x_\mu(s_i)]^2 \right]^{1/2}
\]

If \( s_i \in S \) is presented by symbolic method, then we need another presentation. The order information of a label is reflected though the normalized function of index function. The vagueness information of a label is reflected though the function with the parameter of \( g \).

So \( (O,V) \) is defined as \( \left[ \frac{i}{g} \right] \).

Suppose a preference set \( S^{\mathcal{L}} \) be a linguistic term set with different granularity. Let \( S^g = \{ s_0^g, s_1^g, \ldots, s_n^g \} \) be a linguistic term set. We use two parameters to represent a label’s information.

\[ \text{Definition 5: Linguistic hierarchical is a set of all levels’ linguistic term sets, which is presented as } LH = \cup_{l(t,n(t))} \text{. Each level is denoted as } l(t,n(t)) \text{, where } t \text{ is a number indicating the level of the hierarchy, and } n(t) \text{ is the cardinality of linguistic term set of } t \text{.} \]

So we need to define the new representation model for a linguistic variable in a LH. Let \( S^g = \{ s_0^g, s_1^g, \ldots, s_n^g \} \) be a linguistic term set. We use two parameters to represent a label’s information.

\[ \text{Definition 6: Let } S^g = \{ s_0^g, s_1^g, \ldots, s_n^g \} \text{ be a linguistic term set in hierarchical structure and } R \text{ be a real number set. We define the 2-scale numerical function } 2 \cdot SNF : S^g \to R \text{, which is constituted by two parts: order function and vagueness function } (O,V). \]

\( (O,V) \) should satisfy these conditions:

1) In order to normalize the values of labels in different level, we require \( O \in [0,1] \) and \( V \in [0,1] \);

2) If the linguistic term \( A \) is vaguer than the term \( B \), then \( V(A) > V(B) \).

When linguistic term can be described by fuzzy set or membership function. \( (O,V) \) is defined as \( (x_\mu(s_i), \sigma_\mu(s_i)) \). \( x_\mu(s_i) \) and \( \sigma_\mu(s_i) \) are presented as following by membership function \( \mu_\xi(x) \) [40]:

\[
x_\mu(s_i) = \frac{\int x \mu_\xi(x) dx}{\int \mu_\xi(x) dx}
\]

\[
\sigma_\mu(s_i) = \left[ \frac{\int x^2 \mu_\xi(x) dx}{\int \mu_\xi(x) dx} - [x_\mu(s_i)]^2 \right]^{1/2}
\]

If \( s_i \in S \) is presented by symbolic method, then we need another presentation. The order information of a label is reflected though the normalized function of index function. The vagueness information of a label is reflected though the function with the parameter of \( g \).

So \( (O,V) \) is defined as \( \left[ \frac{i}{g} \right] \).

Suppose a preference set \( S^{\mathcal{L}} \) be a linguistic term set with different granularity. Let \( S^{\mathcal{L}} = \{ s_0^{\mathcal{L}}, s_1^{\mathcal{L}}, \ldots, s_n^{\mathcal{L}} \} \), where \( s_i^{\mathcal{L}} \in S^{\mathcal{L}} \). \( g \) reflects the expert’s uncertainty about a label. For example, there are two drivers describe their driving sense of car A and B, one is
Medium (Low $\cdot$ Medium $\cdot$ High), the other is Medium in (Very Low $\cdot$ Low $\cdot$ Medium $\cdot$ High $\cdot$ Very High). The two labels all reflect that the driving senses of car $A$ and $B$ are almost the same, but the later preference reflects that the driver is more certain about his sense, while the other is more uncertain. If we have more confidence about our preference, then we can subdivide the term set. The confidence is reflected through the granularity. Compared with the previous models, this new definition not only provides a function to present the order in the term set, but also provides a function to present the spread degree which is used to present the term set’s level in a hierarchical structure.

If we suppose the term set to be a symmetrical one with uniform distribution, then we can get the definition 7.

**Definition 7:** Let $S = \{s_0, \cdots, s_g\}$ be a linguistic term set and $R$ be a real number set. We define the $2-SNF: S \rightarrow R$ which is constituted by two parts:

$$O(s_i) = \begin{cases} x(s_i), & \text{when we use extension principle} \\ i / g, & \text{when we use symbolic model} \end{cases} \quad V(s_i) = \begin{cases} 2(x(s_i) - x(s_{i-1})), & \text{when we use extension principle} \\ 2 / g, & \text{when we use symbolic model} \end{cases}$$

There is no difference between the two situations, when the assumed condition of symmetrical and uniform distribution is satisfied. We can prove as following.

**Proof:** $s_i$ is presented by triangular fuzzy number as $(a_i, b_i, c_i)$

$$x_p(s_i) = \frac{\int x u(x) dx}{\int x u(x) dx} = \frac{\int x(b_i - a_i) dx + \int x(c_i - b_i) dx}{\int x(b_i - a_i) dx}$$

$$= \frac{2b_i + a_i + c_i}{4} = b_i = i / g$$

$$= 2(x(s_i) - x(s_{i-1})) = (b_i - b_{i-1}) \times 2$$

$$= (i / g - (i-1) / g) \times 2 / g$$

$$i \in [0, g] \Rightarrow i / g \in [0,1]$$

$$g \in [2, \infty] \Rightarrow 2 / g \in [0,1]$$

Fig. 2 reflects the hierarchical structure of multi-granularity linguistic variables. Fig. 2-a describes the situation where linguistic variables are associated with fuzzy sets, while Fig. 2-b describes the situation where linguistic variables are presented by symbolic model. Firstly, let’s discuss the situation in which two linguistic variables are in the same term set.

Suppose $O(s_i) \in [0,1]$ is the center of $s_i$ which is used to represent the value of label $s_i$. $O(s_i)$ is the most suitable value to describe label $s_i$. So from Fig. 2, we can see the distance between linguistic variables $s_i$ and $s_j$ in the same level can be directly measured by computing the distance of $O(s_i)$ and $O(s_j)$.

$$d(s_i, s_j) = \left\{ \begin{array}{ll} |x(s_i) - x(s_j)|, & \text{extension principle} \\ |i - j| / g, & \text{symbolic model} \end{array} \right.$$
Then let’s give the distance definition of function (17) in which two linguistic variables are in different term sets.

\[
d(s^{(i)}, s^{(j)}) = \sqrt{\left(\sum_{g=1}^{G} \left( (O(s^{(i)}_g) - O(s^{(j)}_g))^2 + (V(s^{(i)}_g) - V(s^{(j)}_g))^2 \right) }^{1/2} + (2 / g_{ij})^2 \], symbolical model (17)
\]

The novel distance measure (16) and (17) should be proven reasonable in theory through satisfying common axioms which have long been taken as a widely acknowledged truth. A metric distance is a function \(d: A \times A \to R\), which satisfies the following three axioms for \(x, y, z \in A\):

(i) \(d(x, y) = 0 \iff x = y\),

(ii) \(d(x, y) = d(y, x)\) (symmetric),

(iii) \(d(x, z) + d(z, y) \geq d(x, y)\) (triangle inequality).

The axiom (i) and (ii) are so obvious that need not to be proved. Now we give the proof of axiom (iii).

\[
d(s^{(i)}, s^{(j)}) + d(s^{(j)}, s^{(k)}) \geq d(s^{(i)}, s^{(k)})
\]

**Proof:** If we plot \(s^{(i)}, s^{(j)}, s^{(k)}\) as coordinates in an order-vagueness plane like Fig. 2-b, we can see \(s^{(i)}, s^{(j)}, s^{(k)}\) are the three points of a triangular. According to the property that sum of two edges is more than the third edge of a triangle. The novel distance measure satisfies the third axioms.

We note that the traditional measurement of variables in multi-granularity linguistic context does not differentiate the semantic difference of variables. In function (17), if two linguistic variables are in the same level of hierarchical structure, then we can conclude \(V(s^{(i)}_g) = V(s^{(j)}_g)\). So the \((V(s^{(i)}_g) - V(s^{(j)}_g))^2\) is neglected in function (16). Function (16) is the special case of function (17), respectively.

4. Application to MADM

In this section, we address this problem under distance measure of multi-granularity linguistic variables. We will discuss the situation where the attribute weights are completely unknown in decision making.

For simplicity, we let \(M = \{1, 2, \ldots, m\}\) and \(N = \{1, 2, \ldots, n\}\). Suppose the alternatives are known. Let \(X = \{x_1, x_2, \ldots, x_n\}\) denote a discrete set of \(n(n \geq 2)\) potential alternatives. Attributes are predefined too. Let \(C = \{c_1, c_2, \ldots, c_m\}\) denote a set of \(m(m \geq 2)\) criterions or attributes. The attribute weights are completely unknown. Let \(D = \{d_1, d_2, \ldots, d_t\}\) be the set of decision makers (DM), and \(\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_t\}\) be the weight vector of DMs, with \(\lambda_k \in [0,1], k = 1, 2, \ldots, t, \sum_{k=1}^{t} \lambda_k = 1\). Suppose that \(A^{(k)} = \{a_j^{(k)}\}_{j=1}^{m}\) is the decision matrix given by the DM \(d_k \in D\), where \(a_j^{(k)}\) is a linguistic variable for alternative \(x_i\) with respect to the attribute \(c_j\).

The standard and mean deviations method is proposed to deal with the uncertain multiple attribute decision making problems, in which the information about the attribute weights are unknown completely [41-43]. The method is proposed based on the idea of maximizing deviations. We improve the method here to determine the optimal relative weights of attributes under linguistic environment.

For decision maker \(d_k\) and the attribute \(c_j\), the deviation of alternative \(x_i\) to all the other alternatives can be expressed as follow:

\[
h^{(k)}(w) = \sum_{j=1}^{m} \sum_{j \neq i} \frac{(a_j^{(k)} - a_i^{(k)})^2}{w_j}
\]

For decision maker \(d_k\) and the attribute \(c_j\), the global deviation of all alternatives can be expressed as follow:

\[
h^{(k)}(w) = \sum_{j=1}^{m} \sum_{j} \sum_{j \neq i} \frac{(a_j^{(k)} - a_i^{(k)})^2}{w_j}
\]

If we consider the problem in a group decision context, we give different weights to decision makers. Then the group global deviation can be expressed as follow:

\[
H(w) = \sum_{k=1}^{t} \lambda_k \sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{j \neq i} \frac{(a_j^{(k)} - a_i^{(k)})^2}{w_j}
\]

For the MADM problems, we need to compare the collective preference values to rank the alternatives. If the performance values of each alternative have little differences under an attribute, it shows that such an attribute plays a small important role in the priority procedure. Contrariwise, if some attribute makes the performance values among all the alternatives have obvious differences, such an attribute plays an important role in choosing the best alternative. Briefly, if the attribute \(c_j\) has little significant difference among alternatives, the
attribute $c_j$ will have little influence on the result of decision. Conversely, the attribute $c_j$ has more influence.

We can get the conclusion: if one attribute has similar attribute values across alternatives, it should be assigned a small weight; otherwise, the attribute which makes larger deviations should be assigned a bigger weight.

Based on the aforementioned analysis, we choose the weight vector $w$ to maximize all deviation values for all the attributes. To present a method to derive the weight vector of attributes in group decision making, we construct a non-linear programming model as follows:

$$\max H(w) = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{m} \sum_{k=1}^{N} \left( d(a_{ij}^{(k)}, a_{ij}^{(k)}) \right)^2 w_j$$

s.t. $w_j \geq 0, j \in M,$

(21)

The weight optimization model is called the maximizing deviation method. The optimization techniques have been successfully applied to classification problems.

The solution to the above optimization problem could be found by MatLab software with optimization toolbox or Lindo/Lingo software package.

In the following, we develop a practical approach based on LWAA operator to decision making with linguistic information.

**Step1.** Identify the evaluation attribute and determine the set of feasible alternatives with the linguistic score for alternatives in terms of each attribute. Construct the linguistic decision matrix $A^{(k)} = [a_{ij}^{(k)}]_{n \times m}$.

**Step2.** Identify the ideal solution which is made of all the best performance scores. $a^*$ is the ideal preference value (label):

$$a^* = \begin{cases} s_+ \text{, attribute is benefit} \\ s_- \text{, attribute is cost} \end{cases}$$

**Step3.** Utilize the maximizing deviation method (21) to obtain the optimal weights.

**Step4.** Ranking all the alternatives $x_i (i \in n)$ and select the best one(s) in accordance with the value of $z_i (i \in n)$. Utilize the LWAA operator to aggregate the decision information. It is based upon the concept that the chosen alternative should have the shortest distance from the ideal solution. The collective overall preference value is $z_i (i \in n)$, which are computed by Eq. (22).

$$z_i = \sum_{k=1}^{r} \sum_{j=1}^{m} w_j d(a_{ij}^{(k)}, a^*)$$

(22)

**Step5.** End.

5. Medical Diagnostic Reasoning via Distances for Multi-Granularity Linguistic Variables

The improvement in fuzzy decision making was successfully verified in a medical diagnostic making system for diagnostic application [44, 45]. In this subsection, we give an example to show how to carry out medical diagnosis by the newly defined distance measure of multi-granularity linguistic variables in Section 3. Different from the example in paper [46], we use linguistic variables to present the characteristic symptoms for the diagnoses considered.

Assume that there are a set of diagnoses $X = \{\text{Viral fever, Malaria, Typhoid, Stomach problem, Chest pain}\}$ (diagnose is presented as $d_k, k = 1, 2, \ldots, 5$), a set of patients $P = \{\text{Al, Bob, Joe, Ted}\}$ (patient is presented as $p_i, i = 1, 2, \ldots, 4$), and a set of symptoms $C = \{\text{Temperature, Headache, Stomach pain, Cough, Chest pain}\}$. The following Table 1 presents the characteristic symptoms for the considered diagnoses. The element $d_{ij}$ in Tab. 1 describes the state of diagnose $d_i$ in terms of symptom $c_j$. And Table 2 gives the symptoms for each patient. The element $p_{ij}$ in Tab. 2 describes the state of patient $p_i$ in terms of symptom $c_j$. One needs to find a proper diagnosis for each patient $p_i$ in the set of patients $P, i = 1, 2, \ldots, 4$. Each element of the tables is given in the form of a linguistic variable, respectively.

Now we consider the medical diagnosis as a pattern recognition problem. We offer personalized medical diagnosis. Our purpose is to distinguish which diagnose the patient’s symptom belongs to. Firstly we have to choose the weight vector $w$ for each patient. Because every patient is different, the weight vector $w$ of patient $p_i$ is different from other. We identify $w$ for each patient according to the symptom’s significant level. If the patient’s symptom $p_{ij}$ has no significant difference among all diagnosis’s symptoms $d_{ij}, p_{ij}$ will have little influence on the result of decision. In order to reduce the influence, the weight of this characteristic symptom should be smaller. The maximizing deviation method is used here to determine $p_i$’s weights.

Then we show how to utilize the proposed method to derive a proper diagnosis for each patient $p_i, i = 1, 2, 3, 4$. We first calculate for each patient $p_i$, the distance measure $d(p_i, d_k)$ between patient symptoms and the set of symptoms that are characteristic for each diagnosis. We construct a non-linear programming model as follows to choose the weight vector $w$ to maximize all deviation values for all the characteristics.
Table 1. Symptoms characteristic $d_{ij}$ for the considered diagnoses.

<table>
<thead>
<tr>
<th></th>
<th>Viral fever</th>
<th>Malaria</th>
<th>Typhoid</th>
<th>Stomach problem</th>
<th>Chest pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>$s_4^8$</td>
<td>$s_8^8$</td>
<td>$s_3^8$</td>
<td>$s_2^8$</td>
<td>$s_0^6$</td>
</tr>
<tr>
<td>Headache</td>
<td>$s_4^8$</td>
<td>$s_8^8$</td>
<td>$s_3^8$</td>
<td>$s_2^6$</td>
<td>$s_0^4$</td>
</tr>
<tr>
<td>Stomach pain</td>
<td>$s_2^8$</td>
<td>$s_0^8$</td>
<td>$s_2^8$</td>
<td>$s_7^8$</td>
<td>$s_2^6$</td>
</tr>
<tr>
<td>Cough</td>
<td>$s_2^4$</td>
<td>$s_9^6$</td>
<td>$s_2^6$</td>
<td>$s_6^4$</td>
<td>$s_1^4$</td>
</tr>
<tr>
<td>Chest pain</td>
<td>$s_5^8$</td>
<td>$s_1^6$</td>
<td>$s_0^4$</td>
<td>$s_2^6$</td>
<td>$s_7^8$</td>
</tr>
</tbody>
</table>

Table 2. Symptoms characteristic $p_{ij}$ for the considered patients.

<table>
<thead>
<tr>
<th></th>
<th>Temperature</th>
<th>Headache</th>
<th>Stomach pain</th>
<th>Cough</th>
<th>Chest pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al</td>
<td>$s_8^8$</td>
<td>$s_5^4$</td>
<td>$s_6^4$</td>
<td>$s_6^6$</td>
<td>$s_0^0$</td>
</tr>
<tr>
<td>Bob</td>
<td>$s_0^6$</td>
<td>$s_3^6$</td>
<td>$s_9^6$</td>
<td>$s_0^6$</td>
<td>$s_0^0$</td>
</tr>
<tr>
<td>Joe</td>
<td>$s_8^6$</td>
<td>$s_8^8$</td>
<td>$s_6^4$</td>
<td>$s_0^4$</td>
<td>$s_0^4$</td>
</tr>
<tr>
<td>Ted</td>
<td>$s_6^8$</td>
<td>$s_4^6$</td>
<td>$s_2^5$</td>
<td>$s_7^8$</td>
<td>$s_2^6$</td>
</tr>
</tbody>
</table>

Table 3. Weights of symptoms for each patient to the considered set of possible diagnoses.

<table>
<thead>
<tr>
<th></th>
<th>Temperature</th>
<th>Headache</th>
<th>Stomach pain</th>
<th>Cough</th>
<th>Chest pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al</td>
<td>0.48</td>
<td>0.09</td>
<td>0.16</td>
<td>0.06</td>
<td>0.21</td>
</tr>
<tr>
<td>Bob</td>
<td>0.21</td>
<td>0.02</td>
<td>0.29</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>Joe</td>
<td>0.32</td>
<td>0.30</td>
<td>0.10</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>Ted</td>
<td>0.29</td>
<td>0.08</td>
<td>0.05</td>
<td>0.43</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Table 4. Similarity for each patient to the considered set of possible diagnoses.

<table>
<thead>
<tr>
<th></th>
<th>Viral fever</th>
<th>Malaria</th>
<th>Typhoid</th>
<th>Stomach problem</th>
<th>Chest problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al</td>
<td>0.24</td>
<td><strong>0.06</strong></td>
<td>0.22</td>
<td>0.46</td>
<td>0.76</td>
</tr>
<tr>
<td>Bob</td>
<td>0.55</td>
<td><strong>0.15</strong></td>
<td>0.21</td>
<td>0.27</td>
<td>0.31</td>
</tr>
<tr>
<td>Joe</td>
<td><strong>0.21</strong></td>
<td>0.44</td>
<td>0.78</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>Ted</td>
<td>0.23</td>
<td>0.41</td>
<td>0.12</td>
<td><strong>0.04</strong></td>
<td>0.22</td>
</tr>
</tbody>
</table>

So we can obtain Table 3 that presents all the weights of different symptoms for the considered patients. Then the proper diagnosis $d_k$ for the $i$th patient $p_i$ is derived according to the similarity measures.

We regard diagnoses as patterns. And we need to find a proper diagnose for each patient. So we utilize the LWAA operator to aggregate the global evaluation of a patient’s symptoms. It is based upon the principle that the patient’s symptoms should have the shortest distance from the proper diagnoses.

Therefore, the collective overall preference values $Z_{ik}$ is computed by Eq. (24).

$$\max H(w^j) = \sum_{i=1}^{5} \sum_{j=1}^{5} (d(p_{ij}, d_{ij}))^2 w^j_k,$$

s.t. $$\left\{ w^j_k \geq 0, \right\}$$

$$\sum_{i=1}^{5} (w^j_k)^2 = 1$$

$$(23)$$

Therefore, we assign to the $i$th patient the diagnosis whose symptoms have the lowest symmetric discrimination information measure from patient’s symptoms. The results for the considered patients are given in Table 4. From Table 4, we can see Al suffers from Malaria, Bob from Malaria, Joe from Viral Fever and Ted from Stomach problem.

6. Conclusion

In MADM, the DMs cannot estimate their preferences with exact numerical values. It is more suitable to provide their preferences by means of linguistic variables rather than numerical ones. The distance measure is an important concept in CWW. Existing methods cannot solve the problems that linguistic variables are with dif-
ferent semantic. We propose a method for calculating distances between multi-granularity linguistic variables and extend the symbolic model. In this article, based on the idea that the attribute with a larger deviation value among alternatives should be evaluated a larger weight, we have developed the maximizing deviation method to determine the optimal relative weights of attributes under linguistic environment. Then we propose a general approach on the basis of LWAA operator to group multi-attribute decision-making problems with linguistic information. The given approach can be applied to medical diagnosis. Dealing with MADM problems defined in multi-granularity linguistic contexts has more application fields which is the future work.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) (71401078, 71171115, 70901043, 71171048 and 71371049), the reform Foundation of Postgraduate Education and Teaching in Jiangsu Province (JGKT10034), Qing Lan Project, Natural Science Foundation of Higher Education of Jiangsu Province of China under Grant (08KJD630002), the Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu (2014SJB059 and 2012SJD630037), and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Reference


**Mei Cai** received the Bachelor of Management degree from Southeast University, Nanjing, China, in 2002, the Master of Management degree from Southeast University, Nanjing, China, in 2005, and the Ph.D. degree from Southeast University, Nanjing, China, in 2012. Her current research interests include decision making under uncertainty, computing with words, and reasoning with uncertainty. She is currently a lecturer of Nanjing University of Information Science and Technology.

**Zaiwu Gong** received his Ph. D. in Management Science and Technology from Nanjing University of Aeronautics and Astronautics. He is currently a Ph.D. tutor. His research interests include group decision making theory, fuzzy sets theory, grey system and evaluation of meteorological disasters risk. He had refereed more than 50 papers for different international journals or international conferences such as Information Science, Soft Computing, etc. He had published more than 70 papers in refereed journals during recent years. He is currently a professor of Nanjing University of Information Science and Technology.

**Jie Cao** received the Ph.D. degree in Management Science and Engineering from Southeast University, Nanjing, China in 2005. He is currently a Professor with the School of Economics and Management, Nanjing University of Information Science and Technology, China. His research interests include Complex System Analysis and Management Decision, Emergency Management, Information Management and Information System, Financial Engineering.

**Minjie Wu** received the Bachelor of Engineering degree from Jiangsu University, Zhenjiang, China, in 2000, the Master of Engineering degree from Southeast University, Nanjing, China, in 2005. Her current research interests include technology innovation in manufacturing industry, supply chain management. She is currently a lecturer of Nanjing University of Information Science and Technology.