Title	A probabilistic model for linguistic multi-expert decision making involving semantic overlapping
Author(s)	Yan, Hong-Bin; Huynh, Van-Nam; Nakamori, Yoshiteru
Citation	Expert Systems with Applications, 38(7): 8901-8912
Issue Date	2011
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/10722
Rights	NOTICE: This is the author's version of a work accepted for publication by Elsevier. Hong-Bin Yan, Van-Nam Huynh, Yoshiteru Nakamori, Expert Systems with Applications, 38(7), 2011, 8901-8912, http://dx.doi.org/10.1016/j.eswa.2011.01.105
Description	

A Probabilistic Model for Linguistic Multi-Expert Decision Making Involving Semantic Overlapping

Hong-Bin Yan[a,*], Van-Nam Huynh[b], Yoshiteru Nakamori[b]

[a]*School of Business, East China University of Science and Technology*
Meilong Road 130, Shanghai, 200237, P.R. China
[b]*School of Knowledge Science, Japan Advanced Institute of Science and Technology*
1-1 Asahidai, Nomi City, Ishikawa, 923-1292, Japan

Abstract

The main objective of linguistic multi-expert decision making (MEDM) is to select the best alternative(s) using linguistic judgements provided by multiple experts. This paper presents a probabilistic model for linguistic MEDM, which is able to deal with semantic overlapping in linguistic aggregation and decision-makers' preference information in choice function. In linguistic aggregation phase, the vagueness of each linguistic judgement is captured by a possibility distribution on a set of linguistic labels. A confidence parameter is also incorporated into the basic model to model experts' confidence degree. The basic idea of this linguistic aggregation is to transform a possibility distribution into its associated probability distribution. The proposed linguistic aggregation results in a set of labels having a probability distribution. As a choice function, a target-oriented ranking method is proposed, which implies that the decision-maker is satisfactory to choose an alternative as the best if its performance is as at least "good" as his requirements. A comparative analysis with prior research is also given to show the advantages of our model via an example borrowed from the literature. The main advantage of our model is its capacity to deal with linguistic labels having partial semantic overlapping as well as incorporate experts and decision-makers' preferences.

Keywords: Decision making; Linguistic assessments; Linguistic aggregation; Semantic overlapping; Experts' confidence; Decision-makers' preferences.

1. Introduction

Multi-expert decision making (MEDM) is a common and important human activity, in which the inherent complexity and uncertainty necessitate the participation of many experts in the decision making process. In practice, the uncertainty, constraints, and even the vague knowledge of the experts imply that the information cannot be assessed precisely in quantitative form, but may be in a qualitative one (Herrera and Martínez,

*Corresponding author
Email addresses: hby19790214@gmail.com (Hong-Bin Yan), huynh@jaist.ac.jp (Van-Nam Huynh), nakamori@jaist.ac.jp (Yoshiteru Nakamori)

2000). A possible way to solve such situation is the use of the fuzzy linguistic approach (Zadeh, 1975a,b,c), which deals with linguistic information that is represented in qualitative terms by means of linguistic variables. Also, the process of activities or decisions usually creates the need for computing with words. As a methodology, computing with words provides a foundation for a computation theory of perceptions or linguistic descriptions (Zadeh, 1999). A key aspect of computing with words is the fusion/aggregation of linguistic variables and computation with vague concepts.

The use of fuzzy sets as the foundation of computing with words is introduced by Zadeh (1975a,b,c). The fuzzy set or membership function associated with each linguistic label is used to represent its semantic. One linguistic computational approach is making use of the associated membership function for each label based on the extension principle, (Degani and Bortolan, 1988; Lin and Wu, 2008, e.g.). Another approach is the symbolic one (Delgado et al., 1993). The idea is that the convex combination of linguistic labels resulting from two linguistic labels should be an element in the set. In these two approaches, however, the results usually do not match any of the initial linguistic labels, hence an approximation process must be developed to express the result in the initial expression domain. This produces the consequent loss of information and lack of precision (Carlsson and Fullér, 2000). To overcome this limitation, a 2-tuple fuzzy linguistic representation model is proposed in Herrera and Martínez (2000). A new approach to extend different classical aggregation operators with the 2-tuple linguistic model is also developed by Herrera and Martínez (2001). Taking a different track, Huynh and Nakamori (2005) have proposed a satisfactory-oriented approach to linguistic MEDM problems by means of the voting mechanism based label semantics proposed by Lawry (2001, 2004). The main idea is that the linguistic MEDM problem is viewed as a decision making under uncertainty (Savage, 1954) framework, and then a probabilistic approach for the pairwise comparison between one alternative and all the others is used to obtain the choice function, i.e., an alternative is the best if its performance is as at least good as all the others. Although such approaches have no loss of information (in some sense) when one applies them in a computational stage for computing with words, they do not directly take into account the underlying vagueness of the linguistic labels, in other words, they assume that any neighboring linguistic labels have no semantic overlapping (Lawry, 2001).

Several approaches have been proposed in an attempt to involve the underlying vagueness of the words in linguistic MEDM problems. In Wang and Hao (2006), Wang and Hao have proposed a new version of 2-tuple fuzzy linguistic representation model based on the symbolic proportion. The main idea of their model is that the experts can express their opinions by not just one label, but spreading that opinion using two adjacent ordinals which are assigned by the experts. Although such approach can model the label overlapping, only two adjacent labels are considered. Also, it is difficult for experts to directly provide precise probability values. Ben-Arieh and Chen (2006a,b) have proposed a fuzzy linguistic OWA (FLOWA) operator, which assigns fuzzy membership functions to all linguistic labels by linearly spreading the weights from the labels to be aggregated. The aggregating result changes from a single label to a fuzzy set with membership levels of each label. Tang (2008) has introduced a collective linguistic MEDM model to capture the underlying vagueness of linguistic labels based on the semantic similarity relation (Tang and Zheng, 2006), in which the similarities among linguistic labels are derived from fuzzy relation of linguistic labels. However, such an approach violates the bounded property of the linguistic aggregation. For more details of the properties of linguistic aggregation, see Delgado et al. (1993). Moreover, it assumes that the same label assessed by different experts has the same label overlapping.

The linguistic judgement provided by one expert implies that the expert makes an assertion. Lawry's *epistemic stance* (Lawry, 2008) implies that when making assertions some things can be correctly asserted while others cannot. Also, the dividing line between those labels are and those that are not appropriate to use may be uncertain, and the assumption that such as division exists would be a natural precursor to any decision making process of the kind just described. Thus when an expert assesses some alternatives (options) with a linguistic label, it is assumed that other linguistic label are appropriate to describe the option. Possibility theory (Dubois et al., 2000) provides a convenient tool to represent experts' uncertain assessments. Furthermore, even if two different experts have assessed an option with the same linguistic label, the appropriateness degree of other linguistic labels may be different according to experts' confidence degree, i.e., to what extent the experts are sure that other linguistic labels are appropriate to describe the option. Finally, our another motivation comes from the fact that experts are not necessarily the decision-makers, but only provide an advice (Shanteau, 2001). For instance, in the *upgrading computer resources* problem in Section 5, there are two agents: a distribution company and four consulting departments. Here, the distribution company acts as the decision-maker and the consulting departments act as the experts. The decision-makers' preference information plays an important role in choice of alternatives, which is missed in most research.

In light of the above observations, we summarize our main contributions as follows. First, we propose a probabilistic approach to linguistic aggregation involving the semantic overlapping. We assume that the appropriate labels are linearly distributed around the linguistic label provided by the expert with a possibility distribution. The label provided by the expert will be called prototype label. Here, possibility has an additional role, namely, that of describing or representing probability distributions. And then based on the basic mass function, we can obtain the probability distribution on the linguistic labels as the aggregation result. Fuzzy modifiers (Zadeh, 1975b) are also used to model some expert' confidence quantifying how he is sure of the appropriateness of other linguistic labels. Our linguistic aggregation approach results in a set of linguistic labels having a probability distribution. Second, we propose a target-oriented ranking method incorporating decision-makers' preferences. We think that the human behavior should be modeled as satisficing instead of optimizing. Intuitively, the satisficing approach has some appealing features because thinking of targets is quite natural in many situations. Third, a thorough comparative analysis is also given to illustrate advantages of our model in terms of four aspects; this point will be manifested in Section 6.

The rest of this paper is organized as follows. Section 2 presents some basic knowledge of linguistic MEDM problems. Section 3 proposes a probabilistic approach to linguistic aggregation involving vague concepts. Section 4 proposes a ranking procedure based on target-oriented decision model, in which decision-makers' preferences are considered. Section 5 provides an illustrative example. Section 6 compares our research with prior approaches in terms of four aspects. Finally, Section 7 presents some concluding remarks.

2. Preliminaries

Linguistic variables provide a means to approximate human activities and human decisions. The concept of a linguistic variable was first introduced in Zadeh (1975a,b,c). In any linguistic approach to solving a linguistic MEDM problem, the label set of a linguistic variable and its associated semantics must be defined first to supply the users with an instrument by which they can naturally express their information. Syntactically, there are two main approaches to generating a linguistic label set. The first one is based on a context-free grammar (Zadeh, 1975a,b,c). This approach may yield an infinite label set. However, according to observations in Miller (1956), the generated language does not have to be infinite. The second approach is to directly supply a finite label set and consider all labels as primary ones, distributed on a scale on which a total order is defined (Bordogna et al., 1997). For our purpose in the following, we only consider linguistic variables with a finite label set. Also, the linguistic label set is totally ordered. For simplicity of notation, we will use an ordered linguistic label set $\mathcal{L} = \{L_0, \ldots, L_n, \ldots, L_N\}$ with $L_0 < \cdots < L_n < \cdots < L_N$ to represent a linguistic variable.

In order to accomplish the objective of choosing the appropriate linguistic descriptors and their semantics for the label set of a linguistic variable, an important aspect need to analyze is the *granularity of uncertainty*, i.e., the level of discrimination among different counts of uncertainty. Typical values of cardinality used in the ordinary linguistic models are odd ones, such as 5, 7, 9, 11, where the middle label presents an assessment of "it may" or "approximately" 0.5, and with the rest of the labels being placed symmetrically around it.

In fuzzy environment, a common characteristic of the MEDM problems, as shown in Table 1, is a finite set of experts, denoted by $\mathcal{E} = \{E_1, \ldots, E_k, \ldots, E_K\}$, who are asked to assess another finite set of alternatives (or options) $\mathcal{A} = \{A^1, \ldots, A^m, \ldots, A^M\}$. The linguistic assessment provided by expert E_k regarding alternative A^m is presented as $x_k^m \in \mathcal{L}$, where $\mathcal{L}$ is a finite, but totally ordered label set of linguistic variables with an odd cardinality, i.e., $\mathcal{L} = \{L_0, \ldots, L_n, \ldots, L_N\}$ with $L_n > L_l$ for $n > l$. Also, each expert is assigned a degree of importance or weight w_k, denoted as $W = [w_1, \ldots, w_k, \ldots, w_K]$.

According to Herrera and Herrera-Viedma (2000), there are mainly three steps to solve this linguistic decision analysis problem:

1. *Choice of the linguistic label set.*
2. *Choice of the aggregation operator for linguistic information.*

Table 1: Linguistic MEDM problem

Alter.	Experts				
	$E_1 : w_1$	$\ldots$	$E_k : w_k$	$\ldots$	$E_K : w_K$
A^1	x_1^1	$\ldots$	x_k^1	$\ldots$	x_K^1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A^m	x_1^m	$\ldots$	x_k^m	$\ldots$	x_K^m
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A^M	x_1^M	$\ldots$	x_k^M	$\ldots$	x_K^M

3. *Choice of the best alternatives*: It is carried out in two phases.
   (a) *Aggregation phase*: Obtaining collective linguistic evaluations of the alternatives by aggregating the individual evaluations by means of the chosen aggregation operator.
   (b) *Exploitation phase*: Establishing a ranking order among the alternatives according to the collective linguistic evaluations.

The first two steps serve the aggregation phase in the third step. From the literature of linguistic decision analysis, there are two general decision models: the first model is mainly based on an aggregation-ranking scheme, and the second one is based on consensus-reaching oriented solution scheme. The model proposed in this paper belongs to the first general class.

3. A Probabilistic Approach to Linguistic Aggregation Involving Semantic Overlapping

In this section we propose a probabilistic approach to linguistic aggregation quantifying an expert's subjective beliefs concerning which labels are appropriate to describe a particular instance (alternative or option). The basic idea of this model is to transform a possibility distribution on linguistic labels into its associated probability distribution based on the mass assignment function.

3.1. Linguistic Aggregation Involving Vague Concepts

Let us return to the linguistic MEDM problem, as shown in Table 1. With the linguistic judgements for alternative A^m provided by a set of experts $\mathcal{E}$, we can obtain a linguistic judgement vector as $X^m = (x_1^m, \ldots, x_k^m, \ldots, x_K^m)$, where $x_k^m \in \mathcal{L}, k = 1, \ldots, K$. When there is no possibility of confusion, we shall drop the subscript m to simplify the notations. Our main objective is to aggregate the linguistic judgement vector X for each alternative A, and then select the best alternatives according to the aggregated results.

The linguistic judgement provided by one expert implies that the expert makes an assertion. It seems undeniable that humans posses some kind of mechanism for deciding whether or not to make certain assertions (e.g., 'The evaluation of a computer system is *high*'). Furthermore, although the underlying concepts are often vague the decision about the assertions are, at a certain level, bivalent. That is to say for an alternative A and a linguistic label L, you are either willing to assert 'A is L' or not. Nonetheless, there seems to be an underlying assumption that some things can be correctly asserted while others cannot. Exactly where the dividing line between those labels are and those that are not appropriate to use may be uncertain, but the assumption that such as division exists would be a natural precursor to any decision making process of the kind just described. This is the main idea of *epistemic stance* proposed by Lawry (2008).

Motivated by the *epistemic stance*, we assume that any neighboring basic linguistic labels have partial semantic overlapping in linguistic MEDM. Thus, when one expert E_k evaluates alternative A using linguistic label $x_k \in \mathcal{L}$, other linguistic labels besides x_k in $\mathcal{L}$ may also be appropriate for describe A, but which of these linguistic labels is uncertain. Here, similar with Lawry and Tang (2009), the linguistic label x_k will be called *prototype label*. Lawry (2004, 2008) introduced a new framework for label semantics where the semantics of linguistic labels are described by appropriateness degrees. The main idea is that the appropriateness measure means the beliefs that linguistic label is appropriate for describing an alternative. If experts can directly assign the appropriateness degrees of all linguistic labels, then we can obtain a possibility distribution. However, the need of experts' involvement creates the burden of decision process. Without additional information, we assume that the appropriate labels are distributed around the prototype label x_k with a linear possibility distribution. Possibility theory is convenient to represent consonant imprecise knowledge (Dubois et al., 2000). The basic notion is the possibility distribution, denoted π. A possibility distribution describes the more or less plausible values of some uncertain variable, i.e., the appropriateness degree of what labels are appropriate for describing an alternative.

It is very rare that when all individuals in a group share the same opinion about the alternatives (options), since a diversity of opinions commonly exists (Ben-Arieh and Chen, 2006b). With the linguistic judgement vector X for alternative A, we can define

$$L_{\min} = \min_{k=1,\ldots,K}\{x_k\} \tag{1a}$$

$$L_{\max} = \max_{k=1,\ldots,K}\{x_k\} \tag{1b}$$

where $x_k \in \mathcal{L}$, $L_{\min} < L_{\max}$, and $L_{\min}$, $L_{\max}$ are the smallest and largest linguistic labels in X, respectively. The label indices of the smallest and largest labels in judgement vector X are expressed as $\texttt{ind}_{\min}$ and $\texttt{ind}_{\max}$, respectively. Also, the label index of the prototype label x_k provided by expert E_k is denoted as $\texttt{pInd}_k$.

Note that, the result of linguistic aggregation should lie between $L_{\min}$ and $L_{\max}$ (including $L_{\min}$ and $L_{\max}$). In addition, if two label indices have the same distance to the index of the prototype label x_k, we assume that they have the same appropriateness (possibility) degree. Furthermore, recall that the linguistic judgement provided by one expert implies that the expert makes an assertion. As Lawry (2008) pointed out,

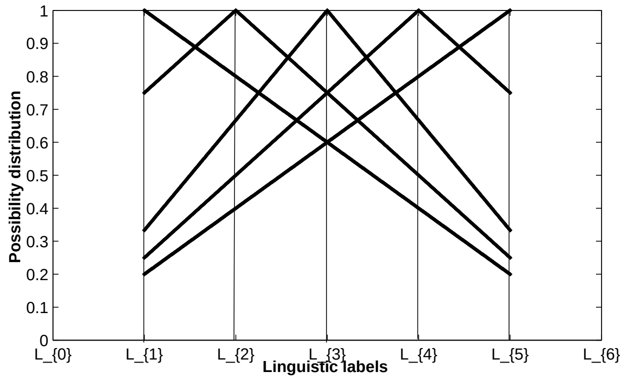

Figure 1: Possible prototype label and its appropriate labels under $[L_1, L_5]$

> An assertability judgement between a 'speaker' and a 'hearer' concerns an assessment on the part of the speaker as to whether a particular utterance could (or is like to) mislead the hearer regarding a proposition about which it is intended to inform him.

Thus if one expert is viewed as a 'speaker', then other experts will act as 'hearer'. Accordingly, we first define a parameter as

$$\Delta_k = \max\{\mathtt{pInd}_k - \mathtt{ind}_{\min}, \mathtt{ind}_{\max} - \mathtt{pInd}_k\}. \tag{2}$$

We then define a possibility distribution of around the prototype label $x_k \in \mathcal{L}$ on linguistic labels L_n as follows

$$\pi(L_n|x_k) = \begin{cases} 1 - \frac{\mathtt{pInd}_k - n}{\Delta_k + 1}, & \text{if } \mathtt{ind}_{\min} \leq n < \mathtt{pInd}_k \\ 1, & \text{if } n = \mathtt{pInd}_k \\ 1 - \frac{n - \mathtt{pInd}_k}{\Delta_k + 1}, & \text{if } \mathtt{pInd}_k < n \leq \mathtt{ind}_{\max} \\ 0, & \text{if } n \notin [\mathtt{ind}_{\min}, \mathtt{ind}_{\max}]. \end{cases} \tag{3}$$

where $n = 0, \ldots, N$. Assume that there is a set of seven linguistic labels $\mathcal{L} = \{L_0, \ldots, L_6\}$. Also, we have $L_{\min} = L_1$ and $L_{\max} = L_5$. Then for a possible prototype label x, according to Eq. (3), we obtain the possibility distribution of appropriate labels as shown in Fig. 1.

The possibility distribution $\pi(L_n|x_k)$ can be viewed as a fuzzy subset. Here, the fuzzy sets play a role of describing or representing probability distributions and the mass assignments provide a semantic for the possibility distributions (Lawry, 2001, 2004). Note $\pi(L_n|x_k)$ is a possibility distribution of around prototype label x_k on the linguistic label set $\mathcal{L}$, then the possibility degrees are reordered as $\{\pi_1(x_k), \ldots, \pi_i(x_k), \ldots, \pi_m(x_k)\}$ such that $1 = \pi_1(x_k) > \pi_2(x_k) > \cdots > \pi_m(x_k) \geq 0$. Then we can derive a consonant mass assignment function $\mathbf{m}_{x_k}$ for the possibility distribution function $\pi(L_n|x_k)$, such

Table 2: Probability distribution on the $N+1$ labels regarding each alternative

Alter.	Linguistic labels				
	L_0	$\dots$	L_n	$\dots$	L_N
A^1	p_0^1	$\dots$	p_n^1	$\dots$	p_N^1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A^m	p_0^m	$\dots$	p_n^m	$\dots$	p_N^m
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A^M	p_0^M	$\dots$	p_n^M	$\dots$	p_N^M

that

$$\begin{aligned}
&\mathbf{m}_{x_k}(\phi) = 1 - \pi_1(x_k)\\
&\mathbf{m}_{x_k}(F_i) = \pi_i(x_k) - \pi_{i+1}(x_k), \text{ for } i = 1, \dots, m-1 \qquad (4)\\
&\mathbf{m}_{x_k}(F_m) = \pi_m(x_k)
\end{aligned}$$

where $F_i = \{\pi(L_n|x_k) \geq \pi_i(x_k)\}, i = 1, \dots, m$ and $\{F_i\}_{i=1}^m$ are referred to as the focal elements of $\mathbf{m}_{x_k}$.

The notion of mass assignment suggests a means of defining probability distribution for any prototype label. Then we can obtain the least prejudiced distribution (Lawry, 2001) of around the prototype label x_k on the linguistic label set $\mathcal{L}$ as follows:

$$p(L_n|x_k) = \sum_{F_i : L_n \in F_i} \frac{\mathbf{m}_{x_k}(F_i)}{|F_i|} \qquad (5)$$

where $L_n \in \mathcal{L}, n = 0, \dots, N$, $\mathbf{m}_{x_k}$ is the mass assignment of $\pi(x_k)$ and $\{F_i\}_i$ is the corresponding set of focal elements. The idea underlying this probability distribution is that, for each focal set F containing linguistic label L, a uniform proportion $\frac{1}{|F|}$ is allocated to L. In other words, the value $p(L_n|x_k)$ reflects the probability that $L_n \in \mathcal{L}$ belongs to the extensions of the prototype label x_k.

Note that each expert is assigned a degree of importance or weight w_k, the weighting vector is denoted as $W = [w_1, \dots, w_k, \dots, w_K]$. The collective probability distribution on the linguistic label set $\mathcal{L}$ is then defined as follows:

$$p_n = p(L_n) = \sum_{k=1}^{K} p(L_n|x_k) \cdot w_k \qquad (6)$$

where $n = 0, \dots, N$. We then obtain a $N+1$-tuple probability distribution on the linguistic label set $\mathcal{L}$ as follows $(p_0, \dots, p_n, \dots, p_N)$ for each alternative A. The probability distributions of all alternatives on the label set $\mathcal{L}$ are shown in Table 2.

Example 1. Assume there are four experts E_1, E_2, E_3, and E_4 to evaluate an alternative A. Each one

Table 3: Possibility distribution and probability distribution on the 7 labels regarding alternative A

Prototype label	$\pi(L)$							$p(L)$						
	L_0	L_1	L_2	L_3	L_4	L_5	L_6	L_0	L_1	L_2	L_3	L_4	L_5	L_6
L_1	0	1	$\frac{3}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	0	0	0	$\frac{25}{48}$	$\frac{13}{48}$	$\frac{7}{48}$	$\frac{3}{48}$	0	0
L_3	0	$\frac{1}{3}$	$\frac{2}{3}$	1	$\frac{2}{3}$	0	0	0	$\frac{3}{36}$	$\frac{7}{36}$	$\frac{19}{36}$	$\frac{7}{36}$	0	0
L_4	0	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	1	0	0	0	$\frac{3}{48}$	$\frac{7}{48}$	$\frac{13}{48}$	$\frac{25}{48}$	0	0

chooses a linguistic label from the set $\mathcal{L}$ to express his opinion. Let us use a linguistic label set as

$$\mathcal{L} = \{L_0, L_1, L_2, L_3, L_4, L_5, L_6\}.$$

Suppose the linguistic evaluations for alternative A provided by the four experts are $X = \{x_1 = L_1, x_2 = L_3, x_3 = L_4, x_4 = L_4\}$. Also, the importance weights of the four experts are $w_1 = 0.25$, $w_2 = 0.25$, $w_3 = 0.25$, and $w_4 = 0.25$, respectively.

Now let us aggregate the linguistic judgement vector using our proposed method. First, we know $L_{\min} = L_1$ and $L_{\max} = L_4$, then according to Eq. (3) we can build the possibility distribution of around prototype label as shown from columns $2-8$ of Table 3. According to the basic mass assignments Eq. (4) and the prejudiced probability distribution Eq. (5), we obtain the corresponding probability distribution on the 7 linguistic labels, as shown from columns $9-15$ of Table 3.

According to Eq. (6) and the weights of the four experts, we can obtain the final probability distribution on the 7 linguistic labels as

$$\left(\frac{0.0}{L_0}, \frac{0.1823}{L_1}, \frac{0.1892}{L_2}, \frac{0.3038}{L_3}, \frac{0.3247}{L_4}, \frac{0.0}{L_5}, \frac{0.0}{L_6}\right).$$

3.2. Incorporating Experts' Confidences into Linguistic Aggregation Involving Vague Concepts

Now we extend the basic model to a general case. In this extended model, we introduce a parameter α to model the confidence/certain degree of an expert. It quantifies to what extent the expert is sure that other linguistic labels around the prototype label are appropriate to describe an alterative. With the confidence character α, we define the possibility distribution of around prototype label $x_k \in \mathcal{L}$ on linguistic label L_n as follows:

$$\pi(L_n|x_k,\alpha) = \begin{cases} \left[1 - \frac{\mathtt{pInd}_k - n}{\Delta_k + 1}\right]^{\alpha}, & \text{if } \mathtt{ind}_{\min} \leq n < \mathtt{pInd}_k \\ 1, & \text{if } n = \mathtt{pInd}_k \\ \left[1 - \frac{n - \mathtt{pInd}_k}{\Delta_k + 1}\right]^{\alpha}, & \text{if } \mathtt{pInd}_k < n \leq \mathtt{ind}_{\max} \\ 0, & \text{if } n \notin [\mathtt{ind}_{\min}, \mathtt{ind}_{\max}]. \end{cases} \tag{7}$$

where α is a linguistic modifier and $\alpha > 0$. When $\alpha > 1$ it means that the expert has an optimistic attitude (he is more sure that the prototype label is appropriate to describe an alternative); when $\alpha = 1$ it means that the expert has a neutral attitude (it is equivalent to the basic model); when $\alpha < 1$ it means that the expert has a pessimistic attitude (he is less sure that the prototype label is appropriate to describe an alternative). Without possibility of confusion, the confidence factor will be also called attitude character. The larger the parameter α is, the more sure the expert is about his assertion, i.e., other labels around the prototype label have small possibilities. Specially, when an expert assigns $\alpha \to +\infty$,

$$\pi(L_n|x_k, \alpha) = \begin{cases} 1, & \text{if } n = \texttt{pInd}_k; \\ 0, & \text{if } n \neq \texttt{pInd}_k. \end{cases} \tag{8}$$

where $\texttt{pInd}_k$ is the label index of the prototype label x_k. It means that the expert E_k is absolutely sure that the prototype label $x_k \in \mathcal{L}$ is appropriate enough to describe alternative A, whereas other linguistic labels are not appropriate to describe A.

Note that each expert can assign different confidence values according to his preferences or belief. In order to better represent expert's attitude factor, we introduce another parameter β, where $\alpha = 2^{\beta}$. Although α and β have continuous forms, for purposes of simplicity, we assign β integer values distributed around 0. For example, $\beta = \{-\infty, \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots, +\infty\}$, consequently we get $\alpha = \{2^{-\infty} \ldots, 1/8, 1/4, 1/2, 1, 2, 4, 8, \ldots, 2^{+\infty}\}$. In order to help experts conveniently express their confidence degree, we construct a totally ordered linguistic label set with an odd cardinality. According to observations in Miller (1956), in practice human beings can reasonably manage to keep about seven labels in mind. We can define the following set of linguistic labels to represent experts' confidence degrees.

$$\begin{aligned} \mathcal{V} = \{&V_0 = \text{absolutely unsure}, V_1 = \text{very unsure}, V_2 = \text{unsure}, \\ &V_3 = \text{neutral}, \\ &V_4 = \text{sure}, V_5 = \text{very sure}, V_6 = \text{absolutely sure}\} \\ \alpha =& \{2^{-M}, 1/4, 1/2, 1, 2, 4, 2^{M}\} \\ \beta =& \{-M, -2, -1, 0, 1, 2, M\} \end{aligned} \tag{9}$$

where M is big enough positive integer to make sure that $[\pi(L_n|x_k)]^{2^M} \to 0$ if $\texttt{ind}_{\min} \leq n < \texttt{pInd}_k$ or $\texttt{pInd}_k < n \leq \texttt{ind}_{\max}$.

And then according to the procedure mentioned in the basic model, Eqs. (4)-(6), we can infer a collective probability distribution for each alternative.

Example 2. We continue to use the aforementioned example. We also assume that the four experts provide four confidence characters, expressed as $\{V_6, V_3, V_1, V_5\}$, the corresponding probability distribution of each expert on the 7 linguistic labels are shown in Table 4.

Table 4: Probability distribution on the 7 labels regarding A with different attitudes

Experts	$p(L)$						
	L_0	L_1	L_2	L_3	L_4	L_5	L_6
E_1	0.0	1.0	0.0	0.0	0.0	0.0	0.0
E_2	0.0	0.0833	0.1944	0.5278	0.1944	0.0	0.0
E_3	0.0	0.1768	0.2214	0.2662	0.3356	0.0	0.0
E_4	0.0	0.0010	0.0205	0.1475	0.8311	0.0	0.0

It is easily seen that, although expert E_3 and E_4 choose linguistic label L_4 as their prototype label, as they have different confidence degrees, thus the derived probability distributions on the linguistic label set $\mathcal{L}$ are different. Expert E_4 has a stronger confidence V_5 than the confidence degree V_1 provided by expert E_3, thus the probability of L_4 derived by E_4 is bigger than that by E_3. Also, we know that expert E_1 has a "absolutely sure" attitude, thus the probability of L_1 is 1.

According to Eq. (6) and the weights of the four experts, we obtain the final probability distribution on the 7 linguistic labels as

$$\left(\frac{0.0}{L_0}, \frac{0.3153}{L_1}, \frac{0.1091}{L_2}, \frac{0.2354}{L_3}, \frac{0.3403}{L_4}, \frac{0.0}{L_5}, \frac{0.0}{L_6}\right).$$

4. Ranking Based on Target-Oriented Decision Model

After linguistic aggregation, the next step of linguistic MEDM is to exploit the best option(s) using a choice function. Most MEDM process is basically aimed at reaching a "consensus", e.g. Bordogna et al. (1997); Herrera-Viedma et al. (2002). Consensus is traditionally meant as a strict and unanimous agreement of all the experts regarding all possible alternatives. The decision model presented below assumes that experts do not have to agree in order to reach a consensus. This assumption is well grounded in research, many of the early decision theories argue that agreement between experts is a necessary condition for expertise. However, experimentation consistently refuted this hypothesis (Einhorn, 1974). An excellent review of this phenomenon of expert disagreement in different domains can be found in Shanteau et al. (2002). There are several explanations that allow for experts not to converge to a uniform opinion. It is well accepted that experts are not necessarily the decision-makers, but provide an advice (Shanteau, 2001). Due to this observation, the linguistic judgements provided by the experts does not represent the decision-makers' preferences.

The inferred probability distribution on a set of linguistic labels for each alternative, as shown in Table 2, could be viewed as a general framework of decision making under uncertainty (Savage, 1954), in which there are $N + 1$ states of nature, whereas the probability distributions are different. In this section, we provide a ranking method considering decision-makers' preferences. Before giving our method, we first briefly recall some basic knowledge of target-oriented decision model.

4.1. Target-Oriented Decision Analysis

The problem of decision making under uncertainty (DUU) (Savage, 1954) that can be most effectively described as follows. $A^m (m = 1, \ldots, M)$ represent the alternatives available to a decision-maker, one of which must be selected. The elements $S_n (n = 1, \ldots, N)$ correspond to the possible values/states associated with the so-called state of nature $\mathcal{S}$. Each element c_n^m of the matrix is the value the decision-maker receives if the alternative A^m is selected and state S_n occurs. The uncertainty associated with this problem is generally a result of the fact that the value of $\mathcal{S}$ is unknown before the decision maker must choose an alternative A^m. Assume a probability distribution $p_{\mathcal{S}}$ on $\mathcal{S} = \{S_1, \ldots, S_n, \ldots, S_N\}$.

As is well-known, the most commonly used method for valuating alternatives A^m to solve the DUU problem is to use the expected utility function:

$$V(A^m) \triangleq \mathtt{EU}(A^m) = \sum_{n=1}^{N} p_{\mathcal{S}}(S_n) U(c_n^m) \tag{10}$$

where U is a utility function defined over universe.

On the other hand, each alternative A^m can be formally considered as a random outcome having the probability distribution p_n defined, with an abuse of notation, as follows:

$$p^m(A^m = x) = p_{\mathcal{S}}(\{S_n : c_n^m = x\}) \tag{11}$$

Then, the target-based model (Bordley and Kirkwood, 2004; Bordley and LiCalzi, 2000) suggests using the following value function

$$\begin{aligned} V(A^m) &\triangleq \Pr(A^m \geq T) \\ &= \sum_x \Pr(x \geq T) \cdot p^m(A^m = x) \\ &= \sum_{n=1}^{N} \Pr(c_n^m \geq T) \cdot p_{\mathcal{S}}(S_n) \end{aligned} \tag{12}$$

where the random target T is stochastically independent of any random value A^m, and $\Pr(x \succeq T)$ is the cumulative distribution function of the target T.

The target-oriented decision model is equivalent to the expected utility model, i.e., probability and utility have a duality. Moreover, target-oriented decision model lies its philosophical root in the bounded rationality (Simon, 1955). As we mentioned previously, the experts are not necessarily the decision-makers, in most cases the decision-makers may have a target in mind. In the sequel we propose a MEDM ranking method based on target-oriented decision model.

4.2. Ranking Based on Target-Oriented Decision Model

Now let us consider the ranking procedure for the probability distribution on $N+1$ linguistic labels in $\mathcal{L}$, as shown in Table 2. We assume that the decision-maker has a target in his mind, denoted as T. We also assume that the target is independent on the set of M alternatives and the linguistic judgements provided by the experts. Based on these assumptions, we define the following function

$$\begin{aligned} V(A^m) &= \Pr(A^m \succeq T) \\ &= \sum_{L \in \mathcal{L}} p^m(A^m = L) \cdot \Pr(L \succeq T) \\ &= \sum_{n=0}^{N} p_n^m \cdot \Pr(L_n \succeq T) \end{aligned} \tag{13}$$

We assume there exists a probability distribution on the uncertain target regarding each linguistic label L_n, denoted as $p_T(L_n)$, where $n = 0, \ldots, N$. Then we define the following function

$$\begin{aligned} V(A^m) &= \Pr(A^m \succeq T) \\ &= \sum_{n=0}^{N} p_n^m \cdot \left[\sum_{l=0}^{N} u(L_n, L_l) p_T(L_l) \right] \end{aligned} \tag{14}$$

Recall that the target-oriented model has only two achievement levels, thus we can define $u(L_n, L_l) = 1$, if $L_n \geq L_l$; 0, otherwise. Then we can induce the following value function

$$\begin{aligned} V(A^m) &= \Pr(A^m \succeq T) \\ &= \sum_{n=0}^{N} p_n^m \cdot \left[\sum_{l=0}^{n} p_T(L_l) \right] \end{aligned} \tag{15}$$

Now let us consider three special cases. Without additional information (if the decision-maker does not assign any target), we can assume that the decision-maker has a uniform probability distribution on the uncertain target T, such that

$$p_T(L_n) = \frac{1}{N+1}, n = 0, \ldots, N. \tag{16}$$

Then we can obtain the value of meeting the uniformly linguistic target as follows:

$$\begin{aligned} V(A^m) &= \sum_{n=0}^{N} p_n^m \cdot \left[\sum_{l=0}^{n} p_{L_l}(T) \right] \\ &= \sum_{n=0}^{N} p_n^m \cdot \frac{n+1}{N+1} \end{aligned} \tag{17}$$

If the decision-maker assigns a specific linguistic label L_l as his target, the probability distribution on uncertain target is expressed as

$$p_T(L_n) = \begin{cases} 1, & \text{if } L_n = L_l; \\ 0, & \text{if } L_n \neq L_l. \end{cases}$$

where $n = 0, \ldots, N$. Then the utility function (probability of meeting target) is as follows:

$$\begin{aligned} V(A^m) &= \Pr(A^m \succeq L_l) \\ &= \sum_{n=0}^{N} p_n^m \cdot \Pr(L_n \succeq L_l) = \sum_{n=l}^{N} p_n^m \end{aligned} \tag{18}$$

In most cases, the linguistic target T specified by the decision-maker is uncertain and it is not so easy for the decision-maker to get the probability distribution of uncertain target on the set of linguistic labels. One possible way to infer the probability distribution of uncertain target is to allow the decision-maker define a a fuzzy target, i.e., a possibility distribution of uncertain target on a discrete interval $[L_{l_1}, L_{l_2}]$, where $L_{l_1}, L_{l_2} \in \mathcal{L}$. And then based on the least prejudiced distribution, we can induce its associated probability distribution function.

Having obtained the utility (probability of meeting target), the choice function for linguistic MEDM model is defined by

$$A^* = \arg \max_{A^m \in \mathcal{A}} \{V(A^m)\} \tag{19}$$

5. Illustrative Example

In this section, we demonstrate the entire process of the probabilistic model via an example borrowed from Herrera and Martínez (2000).

5.1. Problems Descriptions

A distribution company needs to renew/upgrade its computing system, so it contracts a consulting company to carry out a survey of the different possibilities existing on the market, to decide which is the best option for its needs. The options (alternatives) are the following:

A^1	A^2	A^3	A^4
UNIX	WINDOWS-NT	AS/400	VMS

The consulting company has a group of four consultancy departments

E_1	E_2	E_3	E_4
Cost analysis	System analysis	Risk analysis	Technology analysis

Each department in the consulting company provides an evaluation vector expressing its opinions for each alternative. These evaluations are assessed in the set $\mathcal{L}$ of seven linguistic labels, which is expressed as

$$\begin{aligned} \mathcal{L} = \{ & L_0 = \text{none}, L_1 = \text{very low}, L_2 = \text{low}, \\ & L_3 = \text{medium}, \\ & L_4 = \text{high}, L_5 = \text{very high}, L_6 = \text{perfect}\}. \end{aligned}$$

Table 5: Linguistic MEDM problem in upgrading computing resources

Alter.	Experts			
	$E_1 : 0.25$	$E_2 : 0.25$	$E_3 : 0.25$	$E_4 : 0.25$
A^1	L_1	L_3	L_4	L_4
A^2	L_3	L_2	L_1	L_4
A^3	L_3	L_1	L_3	L_2
A^4	L_2	L_4	L_3	L_2

in which $L_n < L_l$ if and only if $n < l$. The evaluation matrix is shown in Table 5. Also, a weighting vector for the four consultancy departments is $W = [0.25, 0.25, 0.25, 0.25]$, i.e., each department is assigned a equal importance.

5.2. Our Proposed Model

As usual, the selection model used to solve this problem consists of two steps:

1. Obtain a *collective performance value* for each alternative.
2. Apply a selection process based on the obtained collective performance vector

In parallel, our proposed model also uses a two-step scheme, but instead:

1. Calculate a group probability distribution for each alternative;
2. Carry out the selection process by the target-oriented model.

In this part, we shall apply our proposed model to solve the above problem.

The first step is to aggregate linguistic assessments involving vague concepts. With the linguistic evaluation matrix (Table 5), we obtain the minimum and maximum linguistic labels for each alternative according to Eq. (1) as follows:

A^1	A^2	A^3	A^4
$[L_1, L_4]$	$[L_1, L_4]$	$[L_1, L_3]$	$[L_2, L_4]$

Without loss of generality, we assume that the four consultant departments have their own confidence degrees. A set of seven linguistic labels, as shown in Eq. (9), is used to represent the consultant departments's confidence degrees. Each consultant department can assign different confidence degrees according to his preference/belief. In this example, we consider four cases:

Case 1: all the four departments assign V_6 = *absolutely sure* as their confidence degrees.

Case 2: all the four departments assign V_3 = *neutral* as their confidence degrees.

Table 6: Probability distributions on linguistic labels with respect to different cases

Cases	Alter.	Linguistic labels						
		L_0	L_1	L_2	L_3	L_4	L_5	L_6
Case 1	A^1	0.0	0.25	0.0	0.25	0.5	0.0	0.0
	A^2	0.0	0.25	0.25	0.25	0.25	0.0	0.0
	A^3	0.0	0.25	0.25	0.5	0.0	0.0	0.0
	A^4	0.0	0.0	0.5	0.25	0.25	0.0	0.0
Case 2	A^1	0.0	0.1823	0.1892	0.3038	0.3247	0.0	0.0
	A^2	0.0	0.2153	0.2847	0.2847	0.2153	0.0	0.0
	A^3	0.0	0.25	0.375	0.375	0.0	0.0	0.0
	A^4	0.0	0.0	0.375	0.375	0.25	0.0	0.0
Case 3	A^1	0.0	0.1089	0.1398	0.2827	0.4685	0.0	0.0
	A^2	0.0	0.3078	0.2478	0.1959	0.2485	0.0	0.0
	A^3	0.0	0.238	0.3832	0.3788	0.0	0.0	0.0
	A^4	0.0	0.0	0.3282	0.4412	0.2306	0.0	0.0
Case 4	A^1	0.0	0.2198	0.2367	0.272	0.2715	0.0	0.0
	A^2	0.0	0.2351	0.2649	0.2649	0.2351	0.0	0.0
	A^3	0.0	0.3021	0.3537	0.3442	0.0	0.0	0.0
	A^4	0.0	0.0	0.3442	0.3537	0.3021	0.0	0.0

Case 3: the four departments assign different values as $\{V_1 = \textit{very unsure}, V_2 = \textit{unsure}, V_5 = \textit{very sure}, V_4 = \textit{sure}\}$ as their confidence degrees, respectively.

Case 4: all the four departments assign $V_1 = \textit{very unsure}$ as their confidence degrees.

According to linguistic aggregation with vague concepts, proposed in Section 3, we obtain different probability distributions for the four alternatives with respect to different cases, as shown in Table 6.

From Table 6, it is easily seen that when the four departments assign a *absolutely sure* attitude, it means that they are absolutely sure that a label L is appropriate for describing an alternative. In this case, the group probability distribution will depend only on the weight information. For instance, for alternative A^2 under case 1, the four departments provide their judgements as $\{L_3, L_2, L_1, L_4\}$ and they have equal weight information, thus the probability distribution on the 7 labels is $(0, 0.25, 0.25, 0.25, 0.25, 0, 0)$.

Having obtained the probability distributions on the 7 linguistic labels with respect to the four cases, now let us rank the four alternatives according to the target-oriented ranking procedure proposed in Section 4.

Table 7: Probability of meeting targets

Cases	Targets	Alternatives			
		A^1	A^2	A^3	A^4
Case 1	T_1	**0.5714**	0.5	0.4643	0.5357
	T_2	**0.5**	0.25	0.0	0.25
Case 2	T_1	0.5387	0.5	0.4464	**0.5536**
	T_2	**0.3247**	0.2153	0.0	0.25
Case 3	T_1	**0.5873**	0.4836	0.4487	0.5575
	T_2	**0.4685**	0.2485	0.0	0.2306
Case 4	T_1	0.5136	0.5	0.4346	**0.5654**
	T_2	0.2715	0.2351	0.0	**0.3021**

In this example, the four consultant departments provide their advice, but do not make decisions. The true decision maker is the distribution company. To renew a computer system, the distribution company may simply looks for the first "satisfactory" option that meets some target. Having this in mind, we first assume that the distribution company does not assign his target. In this case, the distribution company has a uniform target, denoted as T_1. The uniform probability distribution of his target is expressed as

$$\left(\frac{1/7}{L_0}, \frac{1/7}{L_1}, \frac{1/7}{L_2}, \frac{1/7}{L_3}, \frac{1/7}{L_4}, \frac{1/7}{L_5}, \frac{1/7}{L_6}\right).$$

If the distribution company can provide a specific label as his target, for example, the company assigns his target as $T_2 = L_4 = high$, it means that the distribution company is satisfactory to choose an alternative as the best if its performance is at least "good" as *high*.

Table 7 shows the probability of meeting those two targets assigned by the distribution company with respect to four cases of confidence degrees provided by the four consultant departments. From Table 7, option A^4 (VMS) or A^1 (UNIX) is the best choice according to the confidence degrees provided by the four departments and the targets provided by the distribution company. In our linguistic MEDM model, the aggregation phase considers experts' preference information and the choice function considers the decision-makers' preference information.

6. Comparative Analysis with Prior Research

Here, as a comparative analysis, we review four solutions to linguistic MEDM problems (Table 1), namely solution based on two-tuple fuzzy linguistic representation model (Herrera and Martínez, 2000), solution based on satisfactory-oriented principle (Huynh and Nakamori, 2005), solution based on FLOWA (Ben-Arieh and Chen, 2006a,b), and solution based on label similarity semantics (Tang, 2008). All these four

Table 8: Comparisons between our research and related work

Approaches	Features			
	Label overlapping	Bounded property	Preference	
			Expert	Decision-maker
Our model	Yes (Probability distribution)	Yes	Yes	Yes
2-tuple model (Herrera and Martínez, 2000)	No	Yes	No	No
Satisfactory-oriented (Huynh and Nakamori, 2005)	No	Yes	No	No
FLOWA (Ben-Arieh and Chen, 2006a,b)	Yes (Fuzzy membership)	Yes	No	No
Similarity based (Tang, 2008)	Yes (Probability distribution)	No	No	No

approaches are based on the order-based semantics of linguistic labels as well as an aggregation-and-ranking scheme. Table 8 shows the main differences between our model and these four approaches. In the sequel, we compare our model with these four approaches one by one according to the example used in Section 5.

6.1. Solution Based on 2-Tuple representation model

To avoid the loss of information, the 2-tuple fuzzy linguistic representation model was proposed by Herrera and Martínez (2000). In this model, information is represented by means of two-tuple of the form (L, α), where $L \in \mathcal{L}$ and $\alpha \in [-0.5, 0.5)$. Under such a representation, if a value $\beta \in [0, N]$ representing the result of a linguistic aggregation operation, then the two-tuple that expresses the equivalent information of β is obtained by means of the following transformation:

$$\begin{aligned} \triangle\colon [0, N] &\longrightarrow \mathcal{L} \times [-0.5, 0.5) \\ \beta &\longrightarrow (L_n, \alpha) \end{aligned} \tag{20}$$

with $n = \texttt{round}(\beta)$ and $\alpha = \beta - n$, where $\texttt{round}(\cdot)$ is the usual round operator and L_n means the linguistic label having the closest index to β. Inversely, a two-tuple $(L_n, \alpha) \in \mathcal{L} \times [-0.5, 0.5)$ can be equivalently represented by a numerical value in $[0, N]$ by the following transformations:

$$\triangle^{-1}\colon \mathcal{L} \times [-0.5, 0.5) \longrightarrow [0, N] \tag{21}$$

such that $\triangle^{-1}(L_n, \alpha) = n + \alpha$.

When K linguistic information expressed by 2-tuple is available, the aggregation result can be derived by using weighted average operator as follows. Let $x = \{(r_1, \alpha_1), \ldots, (r_k, \alpha_k), \ldots, (r_K, \alpha_K)\}$ be a set of linguistic 2-tuples, the 2-tuple weighted average is computed as

$$\triangle \left(\sum_{k=1}^{K} (r_k + \alpha_k) \cdot w_k \right) \tag{22}$$

The comparison of linguistic information represented by 2-tuples is defined as follows. Let (L_n, α_n) and (L_m, α_m) be two tuples, then:

1. If $n < m$, then $(L_n, \alpha_n) < (L_m, \alpha_m)$;
2. If $n = m$, then:
   (a) If $\alpha_n = \alpha_m$, then $(L_n, \alpha_n) = (L_m, \alpha_m)$;
   (b) If $\alpha_n < \alpha_m$, then $(L_n, \alpha_n) < (L_m, \alpha_m)$;
   (c) If $\alpha_n > \alpha_m$, then $(L_n, \alpha_n) > (L_m, \alpha_m)$.

Applying this approach to the *renewing computing resources* problem, we obtain the collective performance value as

A^1	A^2	A^3	A^4
$(L_3, 0)$	$(L_3, -0.5)$	$(L_2, 0.25)$	$(L_3, -0.25)$

which ranks alternatives in the order $A^1 \succ A^4 \succ A^2 \succ A^3$.

Compared with our model, the 2-tuple based solution does not directly take into account the underlying vagueness of the labels, i.e., it assumes that any neighboring linguistic labels have no semantic overlapping. Moreover, it does not either consider the decision-maker's preference information.

6.2. *Satisfactory-oriented solution*

Huynh and Nakamori (2005) proposed a satisfactory-oriented approach to linguistic MEDM. In their framework, the decision matrix (Table 1) is viewed as a decision making under uncertainty, where the set of experts plays the role of states of the world and the weights of experts play the role of subjective probabilities assigned to the experts. Under such formulation, the problem induces M random preferences, denoted by $X^1, \ldots, X^m, \ldots, X^M$, each X^m for an alternative A^m with associated probability distribution P^m is defined by

$$P^m(X^m = L) = P_E(\{E_k \in E | x_k^m = L\}) \tag{23}$$

for $m = 1, \ldots, M$, $k = 1, \ldots, K$, and $L \in \mathcal{L}$.

They then proposed a choice function defined as follows:

$$\begin{aligned} V(A^m) &= \sum_{n \neq m} \Pr(X^m \geq X^n) \\ &= \sum_{n \neq m} \sum_{L \in \mathcal{L}} \left[P^m(X^m = L) \sum_{\substack{x \in \mathcal{L}, \\ L \geq x}} P^n(X^n = x) \right] \end{aligned} \tag{24}$$

Table 9: Probability distributions on seven linguistic labels: Solution based on Satisfactory-oriented principle

Alter.	Linguistic labels						
	L_0	L_1	L_2	L_3	L_4	L_5	L_6
A^1	0.0	0.25	0.0	0.25	0.5	0.0	0.0
A^2	0.0	0.25	0.25	0.25	0.25	0.0	0.0
A^3	0.0	0.25	0.25	0.5	0.0	0.0	0.0
A^4	0.0	0.0	0.5	0.25	0.25	0.0	0.0

This choice function lies its root in the philosophy of satisfactory-oriented principle, i.e., *it is perfectly satisfactory to select an alternative as the best if its performance is as least "good" as all the others.*

Applying the satisfactory-oriented approach to the *computer updating* problem, we first obtain the probability distribution on the linguistic labels as shown in Table 9. Compared with the probability distribution in Table 6, if all the consulting departments share an *absolutely sure* confidence degree, the aggregation step in our model reduces to Huynh and Nakamori's framework, i.e., the weights of experts play the role of subjective probabilities.

And then based on the satisfactory-oriented based solution, we obtain the final choice function values as

A^1	A^2	A^3	A^4
2.25	1.75	1.4375	1.9375

which ranks alternatives in the order $A^1 \succ A^4 \succ A^2 \succ A^3$.

Although this model acts with the computation solely based on the order-based semantics of the linguistic labels as well as yields the same ranking result in the application example (Huynh and Nakamori, 2005, p. 188), it does not directly take into account the underlying vagueness of the labels. In addition, viewed as a decision making under uncertainty, this model does not consider the decision maker's requirements. In general, in the aggregation step, our model generalizes Huynh and Nakamori's work. In the choice function step, although both our approach and Huynh and Nakamori's work are based on the satisfactory-oriented philosophy, we incorporate decision maker's target preference.

6.3. FLOWA

Ben-Arieh and Chen (2006a,b) proposed a so-called fuzzy linguistic ordered weighted average (FLOWA) aggregation operation. In particular, with the linguistic judgement vector $X = (x_1, \ldots, x_k, \ldots, x_K)$ provided by a set of experts $\mathcal{E}$, there exists a totally ordered linguistic subset, expressed as $\mathcal{R} = \{L_i, \ldots, L_n, \ldots, L_j\}$, where $\mathcal{R} \subseteq \mathcal{L}$, L_i is the smallest label and L_j is the largest one in X. With the weights of experts, $W = [w_1, \ldots, w_k, \ldots, w_K]$, we can obtain a weighting vector $u = [u_i, \ldots, u_n, \ldots, u_j]$ for the linguistic labels in $\mathcal{R}$, where $\sum_{n=i}^{j} u_n = 1$. Note that Ben-Arieh and Chen (2006a,b) viewed the weights of experts as the weights of linguistic labels. This is incorrect as more than one expert may provide the same linguistic label as their judgements.

Then the FLOWA operation is defined as

$$\text{FLOWA}\{L_i, \dots, L_n, \dots, L_j\} = \{(L_n, \mu_{L_n}) | L_n \in \mathcal{R}\} \tag{25}$$

where μ_{L_n} is the fuzzy membership assigned to the n-th linguistic label L_n after aggregating the weights on label sets. The fuzzy membership function is defined as

$$\mu_{L_n} = \sum_{m=0}^{N} \mu_{L_n}^{m}, \tag{26}$$

where $\mu_{L_n}^{m}$ is the membership function of the n-th linguistic label L_n generated from the weighted linguistic label $L_m, L_m \in \mathcal{R}$. Then Ben-Arieh and Chen (2006b) define $\mu_{L_n}^{m}$ as

$$\mu_{L_n}^{m} = \begin{cases} \frac{2(j-n)}{(j-i)(j-i+1)} w_m, & \text{if } m = i; \\ \frac{2(n-i)}{(j-i)(j-i+1)} w_m, & \text{if } m = j; \\ \frac{2(n-i)}{(j-i)(m-i)} w_m, & \text{if } i \leq n \leq m; \\ \frac{2(j-n)}{(j-i)(j-m)} w_m, & \text{if } m \leq n \leq j; \\ 0, & \text{if } m < i \text{ or } m > j. \end{cases} \tag{27}$$

The aggregation result of is normalized, i.e., $\sum_{n=0}^{N} \mu_{L_n} = 1$.

Ben-Arieh and Chen (2006a,b) used Lee and Li's fuzzy mean and standard deviation method (Lee and Li, 1988) to rank the aggregation results, expressed as follows:

$$\overline{x}_{\mu}(A^m) = \sum_{n=0}^{N} n \cdot \mu_{L_n}(A^m) \tag{28a}$$

$$\sigma_{\mu}(A^m) = \left[\sum_{n=0}^{N} n^2 \cdot \mu_{L_n}(A^m) - [\overline{x}_u(A^m)]^2 \right]^{1/2} \tag{28b}$$

where $\overline{x}_{\mu}(A^m)$ and $\sigma_{\mu}(A^m)$ are the fuzzy mean and fuzzy spread of the aggregation result for alternative A^m, respectively. For two alternatives A^m and A^n, the ranking relation is as follows:

1. If $\overline{x}_{\mu}(A^m) > \overline{x}_{\mu}(A^n)$, then $A^m \succ A^n$;
2. If $\overline{x}_{\mu}(A^m) < \overline{x}_{\mu}(A^n)$, then $A^m \prec A^n$;
3. If $\overline{x}_{\mu}(A^m) = \overline{x}_{\mu}(A^n)$, then:
   (a) If $\sigma_{\mu}(A^m) = \sigma_{\mu}(A^n)$, then $A^m \sim A^n$;
   (b) If $\sigma_{\mu}(A^m) < \sigma_{\mu}(A^n)$, then $A^m \succ A^n$;
   (c) If $\sigma_{\mu}(A^m) > \sigma_{\mu}(A^n)$, then $A^m \prec A^n$.

Table 10: Fuzzy memberships of the 7 labels, inferred from the FLOWA

Alter.	Fuzzy memberships						
	L_0	L_1	L_2	L_3	L_4	L_5	L_6
A^1	0.0	0.125	0.25	0.375	0.25	0.0	0.0
A^2	0.0	0.125	0.375	0.375	0.125	0.0	0.0
A^3	0.0	0.1667	0.5	0.3333	0.0	0.0	0.0
A^4	0.0	0.0	0.3333	0.5	0.1667	0.0	0.0

where $\succ$, $\prec$, and $\sim$ represent "*is preferred to*", "*is less preferred to*", and "*indifference*", respectively.

Applying the FLOWA operator to the *computer updating* problem, we obtain the fuzzy membership function for each alternative as shown in Table 10.

Finally, according to ranking criterion, Eq. (28), we obtain the fuzzy mean and fuzzy deviation of the aggregation result as

A^1	A^2	A^3	A^4
$(2.75, 0.9682)$	$(2.5, 0.866)$	$(2.1667, 0.6872)$	$(2.8333, 0.6872)$

which ranks alternatives in the order $A^4 \succ A^1 \succ A^2 \succ A^3$.

This FLOWA model allows the aggregation result with a set of labels with a fuzzy membership function, thus in some sense, it is an linguistic aggregation operation involving the underlying vagueness of linguistic labels. However, there are some drawbacks in this model.

First, there is no theoretical formulation about the FLOWA operation in Ben-Arieh and Chen (2006a,b). Even if the FLOWA operation is rational, the aggregation results should be interpreted as a probability distribution, not a fuzzy membership function. Now let us formulate the FLOWA operator based on our approach. The basic idea of our linguistic aggregation approach is to transform a possibility distribution into its associated probability distribution. According to the judgement vector $X = (x_1, \ldots, x_k, \ldots, x_K)$, there exists a totally ordered linguistic subset, expressed as $\mathcal{R} = \{L_i, \ldots, L_n, \ldots, L_j\}$, where $\mathcal{R} \subseteq \mathcal{L}$, L_i is the smallest label and L_j is the largest one in X. For a prototype label $x_k \in \mathcal{L}$, we can define the possibility distribution of around the prototype label on the the set of linguistic labels as follows

$$\pi(L_n|x_k) = \begin{cases} \frac{n-i}{\mathtt{pInd}_k - i}, & \text{if } i \leq n < \mathtt{pInd}_k \\ 1, & \text{if } n = \mathtt{pInd}_k \\ \frac{j-n}{j-\mathtt{pInd}_k}, & \text{if } \mathtt{pInd}_k < n \leq j \\ 0, & \text{if } n \notin [i, j]. \end{cases}$$

In addition, Yager (2002) has proposed a proportional possibility-probability transformation method, expressed as

$$p(x) = \frac{\pi(x)}{\sum_x \pi(x)}. \tag{29}$$

Using the weight information of experts, we can easily obtain the collective probability distribution on the linguistic labels, which is equivalent to the FLOWA (Ben-Arieh and Chen, 2006a,b). Thus the FLOWA aggregation result should be interpreted as a set of labels with a probability distribution, but not a fuzzy membership function. Our linguistic aggregation approach is making use of mass function, which provides a semantic for the possibility distribution. Thus, the FLOWA aggregation operation is not so rational as ours.

Second, from FLOWA aggregation, Eqs. (26)-(27), we found that if some expert choose L_{i+1} as his prototype label, then the possibility that linguistic label L_i will be chosen is 0. Since L_i and L_{i+2} have the same distance to the prototype label L_{i+1}, we think they should have the same appropriateness degree. Thus FLOWA operator is unsuitable to capture the underlying vagueness of linguistic labels.

Third, the FLOWA aggregation operation cannot represent experts' preference information. Different experts may have different confidence degrees.

Finally, the ranking procedure also does not consider the decision-maker's requirements.

6.4. Solution Based on Semantic Similarity Relation Amongst Linguistic Labels

Recently, Tang (2008) proposed a collective decision model based on the semantic similarities of linguistic labels (Tang and Zheng, 2006) to deal with vague concepts and compound linguistic expressions. The compound linguistic expressions is beyond the scope of our research, thus we only consider the vague concepts in linguistic MEDM problems.

They first defined a similarity relation $< R, \mathcal{L} >$ for a set of basic linguistic labels, expressed as

$$R = [r(L_m, L_n)]_{(N+1)\times(N+1)}, \tag{30}$$

such that $r(L_m, L_n) = r(L_n, L_m) \in [0, 1]$ and $r(L_n, L_n) = 1$ for $m, n = 0, \dots, N$.

Then the probability distribution of L_n on the linguistic label set $\mathcal{L}$ is defined by

$$\forall L \in \mathcal{L},\ p_{L_n,R}(L) = \sum_{F: L\in F} \frac{\mathbf{m}_{L_n}(F)}{|F|} \tag{31}$$

where $\mathbf{m}_{L_n}$ is the consonant mass assignment of L_n and $\{F\}$ is the corresponding set of focal elements.

In the linguistic MEDM problem (Table 1), for alternative A^m, each expert E_k provides a linguistic label x_k^m. The collective probability distribution of linguistic label $x_k^m \in \mathcal{L}$ of a given label L_n on the set of linguistic labels $\mathcal{L}$ is as follows

$$p_n^m = \sum_{k=0}^{K} p_{x_k^m}(L_n) \cdot w_k \tag{32}$$

where $n = 0, \dots, N$ and w_k is the weight of expert E_k.

To rank the alternatives, Tang (2008) suggested two methods expressed as follows:

Table 11: Similarity matrix

Linguistic labels	Similarities						
	L_0	L_1	L_2	L_3	L_4	L_5	L_6
L_0	1.0	0.5	0.0	0.0	0.0	0.0	0.0
L_1	0.5	1.0	0.5	0.0	0.0	0.0	0.0
L_2	0.0	0.5	1.0	0.5	0.0	0.0	0.0
L_3	0.0	0.0	0.5	1.0	0.5	0.0	0.0
L_4	0.0	0.0	0.0	0.5	1.0	0.5	0.0
L_5	0.0	0.0	0.0	0.0	0.5	1.0	0.5
L_6	0.0	0.0	0.0	0.0	0.0	0.5	1.0

Table 12: Probability distribution on the 7 labels inferred from the similarity matrix

Alter.	Probability distribution inferred from similarities						
	L_0	L_1	L_2	L_3	L_4	L_5	L_6
A^1	0.0417	0.1667	0.0833	0.25	0.375	0.0833	0.0
A^2	0.0417	0.2083	0.25	0.25	0.2083	0.0417	0.0
A^3	0.0417	0.2083	0.2917	0.375	0.0833	0.0	0.0
A^4	0.0	0.0833	0.375	0.2917	0.2083	0.0417	0.0

1. Expected value function (Ben-Arieh and Chen, 2006b)

$$V(A^m) = \sum_{n=0}^{N} n \cdot p_n^m \tag{33}$$

2. Pairwise comparison method based on the satisfactory-oriented principle (Huynh and Nakamori, 2005), which is defined as

$$\Pr(A^m > A^l) = \sum_{L \in \mathcal{L}} \sum_{L' > L} p^l(L) p^m(L'). \tag{34}$$

$$A^m > A^l \Leftrightarrow \Pr(A^m > A^l) > \Pr(A^l > A^m). \tag{35}$$

Let us apply the similarity based approach to the *computer updating* problem. The first step in their approach is to define a similarity matrix for the 7 linguistic labels as shown in Table 11.

And then according to Eqs. (31)-(32), we obtain the collective probability distribution on the 7 linguistic labels as aggregation results, as shown in Table 12.

According to the first ranking criterion, Eq. (33), we obtain the expected value for each alternative as

Table 13: Probability distributions with respect to different cases

$\Pr(A^m > A^n)$	A^1	A^2	A^3	A^4
A^1	–	0.5139	0.5903	0.4826
A^2	0.2847	–	0.4392	0.3316
A^3	0.224	0.3316	–	0.2587
A^4	0.3177	0.4392	0.4878	–

A^1	A^2	A^3	A^4
3.0	2.5	2.25	2.75

which ranks alternatives in the order $A^1 \succ A^4 \succ A^2 \succ A^3$.

According to the second ranking criterion, the pairwise comparison matrix is shown in Table 13. Accordingly, the ranking order is $A^1 \succ A^4 \succ A^2 \succ A^3$.

Although Tang's work used a similarity semantic of linguistic labels to represent the label overlapping, there are still some drawbacks.

First, the aggregation result by this model violates the bounded property of aggregation operation, i.e., for a subset $\mathcal{R} = \{L_i, \ldots, L_n, \ldots, L_j\}$, the aggregation result will have nonzero probability distribution on other labels $L_m \notin \mathcal{R}$. For instance, from Table 12, the linguistic aggregation result of option A^1 by this model is a probability distribution having nonzero probabilities in $[L_0, L_5]$. However, in the original evaluation matrix, as shown in Table 5, the assessed labels are L_1, L_3, L_4, thus Tang's approach violates the bounded property of linguistic aggregation.

Second, even two experts assign the same linguistic label as their judgements, the overlapping among linguistic labels may be different. In other words, Tang's approach does not take into account experts' preference information.

Third, Tang's ranking procedure is based on Ben-Arieh and Chen (2006b) and Huynh and Nakamori (2005), thus it does not take into account decision-makers' requirements.

7. Conclusions

In this paper, we have proposed a probabilistic model for MEDM problem under linguistic assessments, which is able to deal with linguistic labels having partial semantic overlapping as well as incorporate experts and decision-makers' preference information. It is well known that linguistic MEDM problems follow a common schema composed of two phases: an *aggregation phase* that combines the individual evaluations to a collective evaluations; and an *exploitation phase* that orders the collective evaluations according to a given criterion, to select the best options. For our model, our linguistic aggregation does not generate a specific linguistic label for each alternative, but a set of labels with a probability distribution, which incorporates experts' vague judgements. Moreover, experts' confidence degree is also incorporated to quantify the appropriateness of linguistic labels other than the prototype label. Having obtained the probability distributions on linguistic labels, we have proposed a target-oriented choice function to establish a ranking ordering among the alternatives. According to this choice function, the decision-maker is satisfactory to select an alternative as the best if its performance is as at least "good" as his requirements.

References

Ben-Arieh, D., Chen, Z., 2006a. Linguistic group decision-making: opinion aggregation and measures of consensus. Fuzzy Optimization and Decision Making 5, 371–386.

Ben-Arieh, D., Chen, Z., 2006b. On linguistic labels aggregation and consensus measure for autocratic decision-making using group recommendations. IEEE Transactions on Systems, Man and Cybernetics–Part A: Systems and Humans 36, 558–568.

Bordley, R., Kirkwood, C., 2004. Multiattribute preference analysis with performance targets. Operations Research 52, 823–835.

Bordley, R., LiCalzi, M., 2000. Decision analysis using targets instead of utility functions. Decisions in Economics and Finance 23, 53–74.

Bordogna, G., Fedrizzi, M., Passi, G., 1997. A linguistic modeling of consensus in group decision making based on OWA operator. IEEE Transactions on Systems, Man and Cybernetics–Part A: Systems and Humans 27, 126–132.

Carlsson, C., Fullér, R., 2000. Benchmarking in linguistic importance weighted aggregations. Fuzzy Sets and Systems 114, 35–41.

Degani, R., Bortolan, G., 1988. The problem of linguistic approximation in clinical decision making. International Journal of Approximate Reasoning 2, 143–162.

Delgado, M., Verdegay, J.L., Vila, M.A., 1993. On aggregation operations of linguistic labels. International Journal of Intelligent Systems 8, 351–370.

Dubois, D., Nguyen, H.T., Prade, H., 2000. Possibility theory, probability and fuzzy sets: Misunderstandings, bridges and gaps, in: Dubois, D., Prade, H. (Eds.), Fundamentals of Fuzzy Sets. Mass: Kluwer, Boston, pp. 343–438.

Einhorn, H.J., 1974. Expert judgement: Some necessary conditions and an example. Journal of Applied Psychology 59, 562–571.

Herrera, F., Herrera-Viedma, E., 2000. Linguistic decision analysis: steps for solving decision problems under linguistic information. Fuzzy Sets and Systems 115, 67–82.

Herrera, F., Martínez, L., 2000. A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Transactions on Fuzzy Systems 8, 746–752.

Herrera, F., Martínez, L., 2001. A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 31, 227–234.

Herrera-Viedma, E., Herrera, F., Chiclana, F., 2002. A consensus model for multiperson decision making with different preference structures. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 32, 394–402.

Huynh, V.N., Nakamori, Y., 2005. A satisfactory-oriented approach to multi-expert decision-making with linguistic assessments. IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics 35, 184–196.

Lawry, J., 2001. A methodology for computing with words. International Journal of Approximate Reasoning 28, 51–89.

Lawry, J., 2004. A framework for linguistic modelling. Artificial Intelligence 155, 1–39.

Lawry, J., 2008. Appropriateness measures: An uncertainty model for vague concepts. Synthese 161, 255–269.

Lawry, J., Tang, Y., 2009. Uncertainty modelling for vague concepts: A prototype theory approach. Artificial Intelligence 173, 1539–1558.

Lee, E.S., Li, R.J., 1988. Comparison of fuzzy numbers based on the probability measure of fuzzy events. Computers & Mathematics with Applications 15, 887–896.

Lin, C.J., Wu, W.W., 2008. A causal analytical method for group decision-making under fuzzy environment. Expert Systems with Applications 34, 205–213.

Miller, G.A., 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, 81–97.

Savage, L.J., 1954. The Foundations of Statistics. Wiley, New York.

Shanteau, J., 2001. What does it mean when experts disagree?, in: Salas, E., Klein, G.A. (Eds.), Linking Expertise and Naturalistic Decision Making. Psychology Press, USA, pp. 229–244.

Shanteau, J., Weiss, D.J., Thomas, R.P., Pounds, J.C., 2002. Performance-based assessment of expertise: How to decide if someone is an expert or not. European Journal of Operational Research 136, 253–263.

Simon, H.A., 1955. A behavioral model of rational choice. The Quarterly Journal of Economics 69, 99–118.

Tang, Y., 2008. A collective decision model involving vague concepts and linguistic expressions. IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics 38, 421–428.

Tang, Y., Zheng, J., 2006. Linguistic modelling based on semantic similarity relation among linguistic labels. Fuzzy Sets and Systems 157, 1662–1673.

Wang, J.H., Hao, J., 2006. A new version of 2-tuple fuzzy linguistic representation model for computing with words. IEEE Transactions on Fuzzy Systems 14, 435–445.

Yager, R.R., 2002. On the instantiation of possibility distributions. Fuzzy Sets and Systems 128, 261–266.

Zadeh, L.A., 1975a. The concept of a linguistic variable and its application to approximate reasoning–Part I. Information Sciences 8, 199–249.

Zadeh, L.A., 1975b. The concept of a linguistic variable and its application to approximate reasoning–Part II. Information Sciences 8, 301–357.

Zadeh, L.A., 1975c. The concept of a linguistic variable and its application to approximate reasoning–Part III. Information Sciences 9, 43–80.

Zadeh, L.A., 1999. From computing with numbers to computing with words–From manipulation of measurements to manipulation of perceptions. IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications 46, 105–119.