Information Sciences 271 (2014) 125–142

Contents lists available at ScienceDirect

Information Sciences

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

Distance and similarity measures for hesitant fuzzy linguistic term sets and their application in multi-criteria decision making

Huchang Liao [a,c], Zeshui Xu [b,*], Xiao-Jun Zeng [c]

[a] Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200052, China
[b] Business School, Sichuan University, Chengdu, Sichuan 610065, China
[c] School of Computer Science, University of Manchester, Manchester M13 9PL, United Kingdom

ARTICLE INFO

Article history:
Received 4 December 2012
Received in revised form 20 February 2014
Accepted 21 February 2014
Available online 1 March 2014

Keywords:
Hesitant fuzzy linguistic term set
Distance measure
Similarity measure
Multi-criteria decision making

ABSTRACT

The hesitant fuzzy linguistic term sets (HFLTSs), which can be used to represent an expert's hesitant preferences when assessing a linguistic variable, increase the flexibility of eliciting and representing linguistic information. The HFLTSs have attracted a lot of attention recently due to their distinguished power and efficiency in representing uncertainty and vagueness within the process of decision making. To enhance and extend the applicability of HFLTSs, this paper investigates and develops different types of distance and similarity measures for HFLTSs. The paper first proposes a family of distance and similarity measures between two HFLTSs. Then a variety of weighted or ordered weighted distance and similarity measures between two collections of HFLTSs are proposed and analyzed for discrete and continuous cases respectively. After that, the application of these measures to multi-criteria decision making problems is given. Based on the proposed distance and similarity measures, the satisfaction degrees for different alternatives are established and are then used to rank alternatives in multi-criteria decision making. Finally a practical example concerning the evaluation of the quality of movies is given to illustrate the applicability and advantage of the proposed approach and the differences between the proposed distance and similarity measures.

1. Introduction

Hesitant fuzzy sets (HFSs), which were first introduced by Torra [30] as an extended form of fuzzy sets, have attracted a lot of attention recently due to their effectiveness and efficiency in representing uncertainty and vagueness [13–16,18,30,46,52]. The motivation for introducing HFSs was that it is sometimes difficult to determine the membership degree of an element to a set, and in some circumstances this difficulty is caused by a doubt between a few different values [30]. Since the HFS permits the membership degree of an element to a given set represented by several possible values between 0 and 1, it can express a decision maker's hesitancy efficiently, especially when two or more sources of vagueness appear simultaneously.

It should be noted that the HFS was introduced to handle the problems that are represented in quantitative situations. In many cases, however, uncertainty is produced by the vagueness of meanings whose nature is qualitative rather than quantitative [4,8,9,24]. For example, when evaluating the "speed" of a car, the linguistic terms such as "fast", "very fast", "slow" may be used; when evaluating the "performance" of a company, the terms such as "good", "medium", and "bad" can be used.

* Corresponding author. Tel.: +86 25 84483382.
E-mail addresses: liaohuchang@163.com (H. Liao), xuzeshui@263.net (Z. Xu), x.zeng@manchester.ac.uk (X.-J. Zeng).

http://dx.doi.org/10.1016/j.ins.2014.02.125
0020-0255/

For such cases, Zadeh [51] proposed the fuzzy linguistic approach, which has been extended into several different models, such as the linguistic model based on type-2 fuzzy sets [31], 2-tuple fuzzy linguistic representation model [10,20,21], the proportional 2-tuple model [32], and so on [6]. However, all these extended models have some serious limitation due to the fact that they assess a linguistic variable by using a single linguistic term rather than following the information provided by decision makers regarding the linguistic variable. Because decision makers may consider several terms at the same time or need a complex linguistic term, such a single linguistic term is often insufficient or very hard to be determined. Motivated by HFSs and linguistic fuzzy sets, Rodríguez et al. [26] proposed the concept of the hesitant fuzzy linguistic term set (HFLTS), which provides a different and more powerful form to represent decision makers' preferences in the decision making process. The HFLTS increases the flexibility and capability of eliciting and representing linguistic information. It permits decision makers to use several linguistic terms to assess a linguistic variable. Thus, it provides many advantages in depicting decision makers' cognitions and preferences.

Rodríguez et al. [26] applied HFLTSs to multi-criteria linguistic decision making problems in which decision makers can provide their assessments by linguistic expressions based on comparative terms, such as "*between very low and medium*", or by simple terms, such as "*very low*; *low*; *medium*; *high*; *very high*". By using HFLTSs and context-free grammar, Rodríguez et al. [27] then proposed a new linguistic group decision making model that facilitates the elicitation of flexible linguistic expressions close to human being's cognitive models for expressing linguistic preferences. Liu and Rodríguez [19] presented a new representation of HFLTSs by means of a fuzzy envelope to carry out the computing with words process. Later, Zhu and Xu [53] introduced the hesitant fuzzy linguistic preference relation (HFLPR) as a tool to collect and represent decision makers' preferences and then investigated the consistency of the HFLPR. In order to apply HFLTSs to solve multi-criteria decision making problems more effectively, we shall pay more attention to the basic characteristics of HFLTSs, in particular distance and similarity measures which are fundamentally important in many scientific fields, such as decision making, pattern recognition, and machine learning [42,43,45,46]. In addition, these measures are also the basis of some well-known methods, such as TOPSIS, VIKOR, ELECTRE. Hence, in this paper, we focus on investigating the distance and similarity measures for HFLTSs, and then apply them to multi-criteria decision making within the context of hesitant fuzzy linguistic circumstances.

Based on this focus, the rest of this paper is organized as follows: Section 2 presents the concepts of the linguistic term sets and hesitant fuzzy linguistic term sets. In Section 3, we first review some known distance and similarity measures and then give the definitions of distance and similarity measures for HFLTSs, based on which several distance and similarity measures for two HFLTSs are introduced. In Section 4, we focus on the distance and similarity measures for two collections of HFLTSs, and establish a variety of weighted distance and similarity measures for discrete and continuous cases respectively. Section 5 gives the application of the proposed distance and similarity measures to multi-criteria decision making. The satisfaction degrees of different alternatives are defined in order to rank alternatives. A practical example concerning the evaluation of the quality of movies is then given to illustrate the applicability and advantage of the proposed approach and the different distance and similarity measures.

2. Linguistic term sets and hesitant fuzzy linguistic term sets

2.1. Linguistic term sets

In the process of decision making, decision makers may feel comfortable and straightforward to provide their knowledge by using linguistic terms that are close to human being's cognitive processes. To model and manage such knowledge with uncertainty, the fuzzy linguistic approach which uses fuzzy set theory to model the linguistic information was proposed by Zadeh in [51]. The linguistic variable, defined as "*a variable whose values are not numbers but words or sentences in a natural or artificial language*", enhances the flexibility and applicability of the decision models and provides good application results in many different fields [25].

Definition 1 [51]. A linguistic variable is characterized by a quintuple $(H, T(H), U, G, M)$, where H is the name of variable; $T(H)$ (or simply T) denotes the term set of H, i.e., the set of its linguistic values, U is a universe of discourse; G is a syntactic rule (which usually takes the form of a grammar) for generating the terms in $T(H)$; and M is a semantic rule for associating each linguistic value X with its meaning, $M(X)$, which is a fuzzy subset of U.

The definition reveals that a linguistic variable is actually established by its linguistic descriptors and semantics. There are different ways to choose the linguistic descriptors and to define their semantics [8,9,25,26,48]. The commonly used approaches for selecting the linguistic descriptors include the *ordered structure approach* and *context-free grammar approach*. The definitions of their semantics can be accomplished in three ways: (1) *semantics based on an ordered structure of the linguistic term set*; (2) *semantics based on membership functions and a semantic rule*; and (3) *mixed semantics*. We pay our attention herein to the *ordered structure approach* and the *semantics based on the ordered structure of the linguistic term set*

By means of supplying directly the term set, the *ordered structure approach* defines the linguistic term set via considering all the terms that are distributed on a scale [8,9,48]. The well-known set of seven linguistic terms is given as: $S = \{s_0 = none,\ s_1 = very\ low,\ s_2 = low,\ s_3 = medium,\ s_4 = high,\ s_5 = very\ high,\ s_6 = perfect\}$. Meanwhile, some other linguistic term sets, which are distributed based on different scales, have been developed as well. For example, Xu [35,37,41] introduced a subscript-symmetric linguistic evaluation scale, which can be defined as follows: