Information Fusion 34 (2017) 43–48

Contents lists available at ScienceDirect

Information Fusion

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

On the syntax and semantics of virtual linguistic terms for information fusion in decision making

Zeshui Xu [a,b,*], Hai Wang [a]

[a] School of Economics and Management, Southeast University, Nanjing, Jiangsu 211189, China
[b] School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China

ARTICLE INFO

Article history:
Received 1 June 2015
Revised 4 June 2016
Accepted 6 June 2016
Available online 7 June 2016

Keywords:
Virtual linguistic terms (VLTs)
Syntax
Semantics
Computing with words

ABSTRACT

The virtual linguistic model is a good technique for linguistic decision making and has been widely used in applications including linguistic information fusion. The main purpose of this paper is to define and specify the syntax and semantics of virtual linguistic terms (VLTs) in detail, and then to serve as the theoretical foundation of the computational models based on VLTs. The syntactical rule generates VLTs by a symbolic transformation, and then the semantic rule presents the semantics of VLTs by means of linguistic modifiers. Based on the syntax and semantics, VLTs could be a possible alternative for solving some current challenges of qualitative information fusion in decision making.

1. Introduction

Methodologies for Computing with Words (CWW) [1] are very useful for decision making problems with qualitative criteria and thus have been widely studied and applied in many practical areas. CWW manipulates natural and artificial linguistic expressions which are less precise than numbers but much closer to human's brain mechanisms. All these linguistic expressions form the domain of possible values of a linguistic variable [2]. To reach a final decision, two scenarios of linguistic decision making models can be considered, which are *from words to words* and *from words to numerical outputs/ranking* [3]. The former outputs a linguistic representation of words, whereas the latter results in a ranking of alternatives based on numerical outputs. Till now, there are several famous linguistic decision making models, such as the membership function-based model [4], the type-2 fuzzy sets-based model [5], the ordinal scales-based model [6], the 2-tuple linguistic model [7] and the virtual linguistic model [8]. From a historical view, the virtual linguistic model can be considered as a variant of the 2-tuple linguistic model. Both of them are very popular as they compute linguistic expressions without loss of information. Moreover, Herrera et al. [3] reported that the 2-tuple linguistic model follows *the from words to words* scenario while the virtual linguistic model falls into the other scenario.

When dealing with linguistic information by a certain computational model, the first and basic step is to choose linguistic term sets (LTSs) with syntax and semantics [9]. Although the virtual linguistic model has been widely applied in information fusion-driven decision making [10,11], its lack of clear representation of syntax and semantics has triggered off some discussions [12]. Recently, Liao *et al.* [13] started the discussion with a special case. When the virtual linguistic terms (VLTs) are balanced and uniformly distributed in the considered domain, they constructed a simple yet meaningful mapping between the VLTs and their semantics graphically. To build a sound foundation of the virtual linguistic model, in this paper, we mainly focus on the investigation of the syntax and semantics of VLTs in a general way. We begin our discussion with a predefined discrete LTS with syntax and semantics. The syntax of a VLT is generated by an algorithm based on proper linguistic modifiers. Then the semantic of a VLT can be derived by modifying the closest original linguistic term to a certain level. Finally, we reconstruct the computational model based on VLTs by some predefined operations following the classical computational models based on the ordered structure.

2. Preliminaries

2.1. Linguistic variables

Given a nonempty domain U, a fuzzy set F on U is characterized by a membership function $\mu_F : U \to [0,1]$. For each u of U, $\mu_F(x)$ represents the membership degree of u in F. Generally, a fuzzy set

* Corresponding author.
E-mail addresses: xuzeshui@263.net (Z. Xu), wanghai17@sina.com (H. Wang).

http://dx.doi.org/10.1016/j.inffus.2016.06.002
1566-2535/

F can be denoted by [14]:

$$F = \int_U \mu_F(u)/u \tag{1}$$

The class of all fuzzy sets on U is denoted by $\mathcal{F}(U)$. Furthermore, given $A, B \in \mathcal{F}(U)$, A is a subset of B, denoted by $A \subseteq B$, which holds if and only if $A(u) \le B(u)$ (for all $u \in U$).

A linguistic variable, whose values are words or sentences in a natural or artificial language, serves as an approximation of characterization of phenomena that is too complex or too ill-defined to be described by a conventional numerical variable. Fuzzy sets are used to represent the restrictions associated with the values of a linguistic variable. The definition of linguistic variable is as follows:

Definition 1. [2]. A linguistic variable is characterized by a quintuple $(X, S(X), U, G, M)$, where X is the name of the variable; $S(X)$ (or simply S) denotes the term set of X with each term being a fuzzy variable denoted generically by s and ranging over the domain U which is associated with the base variable u; G is a syntactic rule for generating the names, s, of values of X; and M is a semantic rule for associating with each s its meaning, $M(s)$, which is a fuzzy set of U.

Remark 1. As suggested by Zadeh [2], three denotations, i.e., the name s, its meaning (semantic) $M(s)$ and its restriction $R(s)$ will be used interchangeably to avoid a profusion of symbols.

A particular s, a name generated by G, is called a term. We can denote it as $S = \{s\}$. An important facet of a linguistic variable is the following two rules:

(1) A syntactic rule, having the form of a grammar, to generate the names of the values of the variable.
(2) A semantic rule, to compute the meaning of each value.

If the number of terms in S is infinite, it is necessary to use an algorithm, rather than a table look-up procedure, to generate the elements of S and compute their semantics.

When generating terms in S, linguistic modifiers play an important role. Given an atomic term, composite terms can be generated by modifying the atomic term to certain levels. Generally, given U, a fuzzy modifier FM on U, is a mapping such that [15]:

$$\begin{aligned} FM &: \mathcal{F}(U) \to \mathcal{F}(U) \\ s &\mapsto FM(s, \delta) \end{aligned} \tag{2}$$

where s is a given term and δ is a real number representing the degree of modification. Two classes of famous modifiers are the power modifiers [16] and the shifting modifiers [17].

2.2. Virtual linguistic model

Generally, a LTS with the semantics defined on the domain U can be denoted by

$$S = \{s_t | t = 0, 1, \ldots, \tau\} \tag{3}$$

where τ is a positive integer. In the computational process of the membership function- based model and the ordinal scales-based model, dealing with these discrete linguistic terms may lead to the loss of information. Thus, Xu [8] extended Eq. (3) to a continuous form $\bar{S} = \{s_\alpha | \alpha \in [0, \tau]\}$. Given $s_\alpha \in \bar{S}$, if $s_\alpha \in S$, then it is an original linguistic term (atomic term); otherwise, it is a VLT. Due to the lack of syntax and semantics, Xu [8] had to state that the VLTs can only appear in operations. Based on some simple operational laws, the virtual linguistic model is convenient for information fusion for decision making without loss of information. However, the output VLTs limit the interpretability of this kind of decision making methods.

3. The syntax and semantics of VLTs

In order to generate VLTs, it is natural to begin with a predefined LTS associated with semantics having the form of Eq. (3). Note that the symbol s_t is used to represent both the name of the term and its semantic taking the form of a fuzzy set defined on the domain U.

Each linguistic term $s_t \in S$ is called an original linguistic term and considered as an atomic term. VLTs are generated by the proper linguistic modifiers based on the original linguistic terms. Roughly, we will generate a VLT by its closest original linguistic term. It is easy to generate a new VLT by a symbolic transformation as follows:

Definition 2. (Syntactical generation of a VLT). Let $S = \{s_t | t = 0, 1, \ldots, \tau\}$ be a LTS with the semantics defined on the domain U. For any $t \in \{0, 1, \ldots, \tau\}$, let

$$\delta \in \begin{cases} [0, 0.5), & t = 0 \\ [-0.5, 0], & t = \tau \\ [-0.5, 0.5), & else \end{cases} \tag{4}$$

then the pair (t, δ) generates a VLT s_α, with $\alpha = t + \delta$. The set of VLTs is denoted by $\bar{S} = \{s_\alpha | \alpha \in [0, \tau]\}$.

According to Definition 2, a VLT, s_α, is generated by an atomic term s_t and a real number δ satisfying $t = round(\alpha)$ and $\delta = \alpha - t$, where *round* is the classical round function. The original linguistic term can be viewed as a special VLT with $\delta = 0$.

Example 1. Given the following LTS (whose semantics are shown in Fig. 1):

$$\begin{aligned} S = \{ & s_0 = extremely\ poor,\ s_1 = very\ poor,\ s_2 = poor, \\ & s_3 = slightly\ poor,\ s_4 = fair,\ s_5 = slightly\ good, \\ & s_6 = good,\ s_7 = very\ good,\ s_8 = extremely\ good\} \end{aligned}$$

Let $t = 5$ and $\delta = 0.4$, a new VLT, named $s_{5.4}$, can be generated. Note that, different from original linguistic terms, it is hard to name a VLT by words or sentences exactly. For example, s_5 can be named by *slightly good*, but $s_{5.4}$ can not be endowed with a linguistic name.

Till now, only the symbolic has been generated for a VLT. Their meanings and semantics should be assigned. This will be completed by linguistic modifiers defined below:

Definition 3. (The semantic rule). Let $S = \{s_t | t = 0, 1, \ldots, \tau\}$ be a LTS with the semantics defined on the domain U. For any $\alpha \in [0, \tau]$, the semantic of the VLT s_α generated by Definition 2 is given by

$$s_\alpha = FM(s_t, \delta) \tag{5}$$

where $t = round(\alpha)$, $\delta = \alpha - t$ and FM is a linguistic modifier on U.

Remark 2. This definition only supplies a strategy to obtain the semantics of VLTs. The linguistic modifier should be determined according to the type of original terms distributed on the domain. We will specify the choice of linguistic modifiers in the coming subsections.

Remark 3. Similar to the 2-tuple linguistic model [7], the syntax and semantic of a VLT should be clarified by two parameters. There is a pair of symbolic transformation functions in 2-tuple linguistic model. One is used to transform a 2-tuple linguistic term to a real number $\beta \in [0, \tau]$ which represents the equivalent linguistic information and the other is used for inverse transformation. However, there are some different details. In the syntactical aspect, a 2-tuple linguistic term is generated by a real number β, whereas a VLT