Information Sciences 181 (2011) 4394–4408

Contents lists available at ScienceDirect

Information Sciences

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

A hierarchical model of a linguistic variable

Adolfo R. de Soto [1]

School of Industrial Engineering and Computer Science, University of León, 24071 LEON, Spain

ARTICLE INFO

Article history:
Available online 12 April 2011

Keywords:
Interpretable fuzzy systems
Computing with words
Linguistic variables
Multiresolution analysis
Hierarchical models
Fuzzy sets

ABSTRACT

In this work a theoretical hierarchical model of dichotomous linguistic variables is presented. The model incorporates certain features of the approximate reasoning approach and others of the Fuzzy Control approach to Fuzzy Linguistic Variables. It allows sharing of the same hierarchical structure between the syntactic definition of a linguistic variable and its semantic definition given by fuzzy sets. This fact makes it easier to build symbolic operations between linguistic terms with a better grounded semantic interpretation. Moreover, the family of fuzzy sets which gives the semantics of each linguistic term constitutes a multiresolution system, and thanks to that any fuzzy set can be represented in terms of the set of linguistic terms. The model can also be considered a general framework for building more interpretable fuzzy linguistic variables with a high capacity of accuracy, which could be used to build more interpretable Fuzzy Rule Based Systems (FRBS).

1. Introduction

The concept of Fuzzy Linguistic Variable is a staple of the Fuzzy Set Theory [41,48]. It has the remarkable property of putting together symbols and the meaning of those symbols as proper elements of a computational system [47]. However, calculating with both levels simultaneously is not a trivial problem. Many symbolic operations have a very large set of possible interpretations as fuzzy operations. Besides that, many fuzzy operations on fuzzy sets do not have a clear correspondence with symbolic operations or do not correspond to inner operations in the set of symbolic terms. For these reasons, both aspects, the symbolic and the aspect relative to the meaning, have been developed differently in Fuzzy Set Theory and its applications. Calculus with fuzzy sets has had a higher prominence than calculus with symbols or words. The main hypothesis of this work is that a stronger structure within the set of linguistic terms of a linguistic variable is needed, but with the condition of a semantic model being associated with the same structure. In this way, inner symbolic operations would also have as a result a fuzzy set which gives their meaning.

Essentially, the linguistic variable concept introduces two levels for manipulating "words", the syntactic or symbolic level, where the names of the words are given and certain operations can be defined working on those symbols to generate new symbols, and the semantic or meaning level, where Fuzzy Sets are introduced to give the meaning of each symbolic word. Both levels are expressed explicitly in the initial definition of a Linguistic Variable given by Zadeh [43–45]. In the original definition, no structure is considered for the set of linguistic terms, in spite of the fact that both linguistic connectives such as *and, or, not* and linguistic modifiers such as *very, little, at most,* were considered from the beginning for building derived terms from the basic terms. Zadeh's definition does not fix either the semantic mapping which gives the semantic value of each linguistic term. Traditionally two different ways have been considered. The first consists of defining a different fuzzy set for each linguistic term, regardless of whether the term is simple or includes various operators or modifiers. This approach has usually been used in Fuzzy Rule Based Systems. In contrast, the second approach gives a meaning to the basic

[1] This work has been partially funded by the Spanish Ministry of Science and Innovation through Project TIN2008-06890-C02-01 and the Human Resources National Mobility Program through grant PR2008-0143.
E-mail address: adolfo.rdesoto@unileon.es

0020-0255/$ - see front matter
doi:10.1016/j.ins.2011.04.006

terms by means of fuzzy sets directly but calculates the meaning of the derived linguistic terms. Using operations on the class of fuzzy sets, a meaning for the connectives and for the linguistic modifiers is set by means of certain fuzzy operations and then those operations are used to carry out the calculation. As is well known, t-norms, t-conorms [4] and negation functions [36] are used to represent connectives, while several types of fuzzy operators are employed to represent linguistic modifiers [6,31,32].

In this work the first approach described above is considered and a hierarchical interpretation of a dichotomous Fuzzy Linguistic Variable is given. This kind of variable is characterized by the use of two antonymous linguistic terms as basic terms. The use of this type of linguistic variable is very common in natural language and human beings process a great amount of information using linguistic terms derived from these basic antonymous linguistic terms. The use of this kind of variable is perhaps an embodied property of the human mind [26], and its representation as a hierarchical classification system is able to show the two basic properties of hierarchical classifications [34]: cognitive economy and correspondence with the structure of the world, this latter being understood as the pairing of world and observer. Furthermore, in this type of linguistic variable the middle term is very important. In the model presented here, the middle term comes from one assumption in the model.

The linguistic variables used in Fuzzy Rules Based Systems customarily have a finite and totally ordered number of linguistic terms. It is evident that with a finite number of linguistic terms the capacity for accuracy is restricted. In fact, when more accuracy is needed in a fuzzy rule based system the usual procedure is to increase the number of fuzzy labels. Working with a finite number of labels is easier than using an infinite number. For that reason, many expert systems use a finite number of symbolic linguistic terms, without a fuzzy interpretation, to carry out reasoning, following the model of finite multivalued logics. The problem with this approach is that a serious restriction is imposed on the accuracy.

On the contrary, if a natural language is considered, an infinite number of linguistic terms seems more plausible. At least potentially, natural languages have the possibility to produce a non-bounded number of linguistic terms associated with a linguistic variable through the use of connectives or linguistic modifiers. As far as approximate reasoning is concerned, this would appear to be true. Nevertheless, it is also true that human beings make use of only a few linguistic terms, and employ linguistic modifiers to build new linguistic expressions through the intensification or weakening of the meaning of a linguistic term. In this way, it is possible to attain greater accuracy. However, the number of different linguistic expressions that people use in their vocabulary is quite limited, in spite of their potential capacity to produce an infinite number of them.

The model in this work is also based on a multiresolution scheme which permits working simultaneously with several levels of resolution, and in this way tries to solve the problem of finite resolution *versus* potentially infinite resolution. In many problems that human beings solve easily, resolution is set at the level which is sufficient to solve the problem in question. For example, to find out if a piece of furniture will fit against a wall, people use an accuracy of centimetres, which is good enough. On the contrary, for a carpenter to make that piece of furniture probably requires an accuracy of millimetres. The objective is to be able to work in Fuzzy Rule Based Systems with an adequate resolution for the problem. Multiresolution Analysis Theory, which was introduced by Mallat in [27], is usually used in Wavelet Theory [9] as the main method for building wavelet functions.

Another important issue in this work is the matter of interpretability of Fuzzy Rule Based Systems [7]. At the beginning of their development, human interpretability of FRBS was perceived to be an essential aspect of them. However, in many works on the topic of FRBS accuracy is the main concern. This work considers interpretability as a fundamental property of FRBS and interpretability is reached through a simple structure in the linguistic label set and through having as few rules as possible. Nevertheless, the problem is how to achieve a high degree of accuracy with few rules and a simple linguistic term set. A hierarchical model can help to solve this problem. Hierarchical Fuzzy Rule Based Systems have frequently been used for modeling systems [35,39] but fewer examples exist that use Hierarchical Linguistic Variables [14,19].

The approach to interpretability of FRBS adopted here is mainly focused on getting a good structure in the set of linguistic terms of a Linguistic Variable used in FRBS. Following the taxonomy given in [18], this paper is only concerned with the semantic interpretability of a Linguistic Variable. But a relevant difference with usual approaches to FRBS interpretability (see, for example, [2,3,7,22,23]) has to be remarked. The final goal here is not to obtain a high evaluation in the interpretability of one FRBS in an isolated way, but to build interpretable Linguistic Variables to be used in several related FRBS which incorporate a common variable with different levels of precision. For example, in the field of home automation, several FRBS can use the temperature variable with different precision in different systems. Using the same set of linguistic terms for the temperature, in spite of the different precision required in each FRBS, can improve the interpretability of the whole system.

Several other works have studied the problem of the interpretability in Fuzzy Rule Based Systems through defining an adequate structure in the set of linguistic terms of a Linguistic Variable. In their works, Cat Ho and collaborators [8] built an abstract algebraical structure to model linguistic modifiers as operations on the set of linguistic terms. This was solely a symbolic approach and the meaning of the terms was not considered. Work reference [33] modeled linguistic expressions of natural language which involved trichotomous linguistic variables, i.e. linguistic variables with three basic linguistic terms, such as *small–medium–large* using a formal system of higher-order Fuzzy Logic. Again, being a symbolic approach, this model was independent of the meaning given to the basic linguistic terms and to the linguistic modifiers. Finally, work reference [5] introduced a definition of interpretability in Linguistic Variables by means of defining certain relationships within the set of linguistic terms which are preserved in the semantic interpretation counterpart (given by fuzzy sets). None of these works considers the approach adopted here, where the meaning of every linguistic term is defined directly without using operators to derive the meaning of a composite linguistic term. Similarly, a multiresolution perspective has not been