Information Fusion 41 (2018) 176–181

Contents lists available at ScienceDirect

Information Fusion

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

# A new proposal to deal with hesitant linguistic expressions on preference assessments

Marie Durand [a,b], Isis Truck [b,*]

[a] UMR SFL 7023, Paris 8 University, 2 rue de la Liberte, Saint-Denis F-93526, France
[b] CHArt Laboratory – EA 4004, Paris 8 University, 2 rue de la Liberte, Saint-Denis F-93526, France

ARTICLE INFO

*Article history:*
Received 14 April 2017
Revised 29 July 2017
Accepted 10 September 2017
Available online 14 September 2017

ABSTRACT

Information fusion and hesitant information fusion represent an important part of decision making processes. This paper focuses on hesitant expressions and the way to take them into account in the computations, using weights served by a simple but efficient process. In a previous paper we have proposed to use an operator called the *symbolic weighted median* to express hesitant linguistic assessments such as "I hesitate between this and that but I tend to lean toward that alternative". Now we go further in explaining in detail how to transform such expressions into our hesitant operators. Inspired by language science research, several hesitant linguistic expressions are discussed, including linguistic modifiers and qualifiers, then they are transformed into weight vectors before being aggregated to complete information fusion.



## 1. Introduction

Hesitation between several alternatives is very common in decision making. Therefore there is a need for dealing with uncertainty during fusion information to make a decision. A number of works and studies have been carried out and propose tools such as a hesitant fuzzy linguistic framework [1–6] and some of these tools use the 2-tuple linguistic model [7]. The basic concept is the following: people may hesitate between several alternatives while an alternative is expressed through a fuzzy set or a *linguistic fuzzy 2-tuple* such as a pair $(s_i, \alpha)$ where $s_i$ is a linguistic term and $\alpha$ a number that represents a symbolic translation to avoid loss of information during the computations. There is a real need in practical cases, especially complex decision making problems [8]. Indeed, for many real-world decisions, the knowledge is either absent, or may only be known in some vague, hesitant, intuitive, way [9]. Of course, other models of 2-tuples can also be considered and we proposed in a previous paper a classification of several hesitant operators using several 2-tuple models [10].

However, a recent paper pointed out the low quality of some proposals and discussed which direction new proposals on hesitant fuzzy sets should follow [11]. Indeed, one of the challenge is: "How can we represent human knowledge?" and the advice given is: "future extensions must be discussed in the context of representing uncertainty in a real world context, providing useful tools for those problems that require representing and managing the hesitancy in expert's knowledge." That is why our purpose here is to continue our work about the transformation of hesitant linguistic expressions into hesitant operators [10]. This is not only a way to represent human knowledge, but a way to represent uncertainty in a real world context because linguistic expressions come from subjective assessments, judgements, feelings... from databases to study the mechanisms that underlie second language acquisition [12,13]. Another question arises: what can be learnt from the language science research?

In this recent paper, our proposal was to use an aggregation operator called the *symbolic weighted median* to express hesitant assessments such as "I hesitate between $\tau_2$ and $\tau_3$ but I tend to lean toward $\tau_3$", where $\tau_i, i \in \{0, n-1\}$ is one among $n$ alternatives. How to go from the linguistic expression to the mathematical modelling was future work. Now we are interested in this question and in the problem of the linguistic expressions themselves.

The present paper is organized as follows: Section 2 recalls the *Symbolic Weighted Median* and their underlying tools, the *Generalized Symbolic Modifiers*. The third section shows the importance of the weights according to the linguistic expression of hesitationthat is divided into *qualifiers* and *modifiers*. Section 4 details the way to obtain those weightswith a mapping function while Section 5 defines a formal method to aggregate such linguistic hesitant expressions. Finally Section 6 concludes this study.

* Corresponding author.
*E-mail addresses:* marie.durand@cnrs.fr (M. Durand), i.truck@gmail.com, isis.truck@univ-paris8.fr (I. Truck).

http://dx.doi.org/10.1016/j.inffus.2017.09.007
1566-2535/

**Table 1**
Summary of reinforcing and weakening GSMs, according their nature [14].

| MODE NATURE | Weakening | | Reinforcing | |
|---|---|---|---|---|
| Erosion | $\tau'_j = \tau_{\max(0,i-\rho)}$ $\mathcal{L}_{M'} = \mathcal{L}_{\max(1,M-\rho)}$ | ER($\rho$) | $\tau'_j = \tau_i$ $\mathcal{L}_{M'} = \mathcal{L}_{\max(i+1,M-\rho)}$ | EW($\rho$) |
| | | | $\tau'_j = \tau_{\min(i+\rho,M-\rho-1)}$ $\mathcal{L}_{M'} = \mathcal{L}_{\max(1,M-\rho)}$ | ER′($\rho$) |
| Dilation | $\tau'_j = \tau_i$ $\mathcal{L}_{M'} = \mathcal{L}_{M+\rho}$ | DW($\rho$) | $\tau'_j = \tau_{i+\rho}$ $\mathcal{L}_{M'} = \mathcal{L}_{M+\rho}$ | DR($\rho$) |
| | $\tau'_j = \tau_{\max(0,i-\rho)}$ $\mathcal{L}_{M'} = \mathcal{L}_{M+\rho}$ | DW′($\rho$) | | |
| Conservation | $\tau'_j = \tau_{\max(0,i-\rho)}$ $\mathcal{L}_{M'} = \mathcal{L}_M$ | CW($\rho$) | $\tau'_j = \tau_{\min(i+\rho,M-1)}$ $\mathcal{L}_{M'} = \mathcal{L}_M$ | CR($\rho$) |

## 2. Preliminaries

### 2.1. Previous works on modifiers and median

The *Generalized Symbolic Modifiers* (GSMs) have been proposed in [14] and are used thereafter to express the result of information fusion through an operator called the *Symbolic Weighted Median* (SWM) that is entirely defined by GSMs. A GSM is associated to a semantic triplet of parameters: radius (denoted $\rho$ — the more the radius, the more powerful the modifier), nature (i.e. dilated, eroded or conserved) and mode (i.e., reinforcing, weakening or centring). GSMs are defined through a totally ordered set of $M$ alternatives $\mathcal{L}_M = \{\tau_0, \dots, \tau_i, \dots, \tau_{M-1}\}$ $(\forall i, j \in \{0, 1, \dots, M-1\}, \tau_i \le \tau_j \Leftrightarrow i \le j)$. Four basic operators are defined $\vee$ (max), $\wedge$ (min), $\neg$ (symbolic negation, with $\neg\tau_j = \tau_{M-j-1}$) and the Łukasiewicz implication $\to_L$: $\tau_i \to_L \tau_j = \min(\tau_{M-1}, \tau_{M-1-(i-j)})$.

$\tau'$, the value after modification, is computed according to a GSM $m$ with a radius $\rho$, denoted $m_\rho$. Actually $m_\rho$ modifies the pair $(\tau_i, \mathcal{L}_M)$ into another pair $(\tau'_j, \mathcal{L}_{M'})$.

**Definition 1.** [14]
Given $\rho \in \mathbb{N}^*, i \in \{0, \dots, M-1\}$, any $\tau'_j (j \in \{0, \dots, M'-1\})$ can be computed through $m_\rho(\tau_i)$.

$$m_\rho : \mathcal{L}_M \to \mathcal{L}_{M'}$$
$$\tau_i \mapsto \tau'_j$$

Three GSM families have been defined: weakening, reinforcing (see Table 1) and central ones (see Definition 2 for an example of such a GSM, where *DC′* is a *d*ilated *c*entring modifier, *i.e.* granularity increases).

**Definition 2.** [14]

$$\mathrm{DC}'(\rho) = \begin{cases} \tau'_j = \begin{cases} \tau_{\frac{i*(M*\rho-1)}{M-1}} & \text{if } \tau_{\frac{i*(M*\rho-1)}{M-1}} \in \mathcal{L}_{M'} \\ \tau_{\lfloor \frac{i*(M*\rho-1)}{M-1} \rfloor} & \text{otherwise (pessimistic)} \\ \tau_{\lfloor \frac{i*(M*\rho-1)}{M-1} \rfloor + 1} & \text{otherwise (optimistic)} \end{cases} \\ \mathcal{L}_{M'} = \mathcal{L}_{M*\rho} \end{cases}$$

**Definition 3.** [15] Let $\mathcal{L}_M = \{\tau_0, \tau_1, \dots, \tau_{M-1}\}$ be a collection of $M$ ordered elements. When the elements have weights, the collection is denoted $\langle \tau_0^{w_0}, \tau_1^{w_1}, \dots, \tau_{M-1}^{w_{M-1}} \rangle \in \mathcal{B}^{\mathcal{L}_M}$ (set of collections) such that $\sum w_i = 1,\ i = \{0, \dots, M-1\}$. The *Symbolic Weighted Median* $\mathcal{M}$ is defined as follows:

$$\mathcal{M}: \quad \mathcal{B}^{\mathcal{L}_M} \to \mathcal{L}_{M'}$$
$$\langle \tau_0^{w_0}, \tau_1^{w_1}, \dots, \tau_{M-1}^{w_{M-1}} \rangle \mapsto \mathcal{M}(\langle \tau_0^{w_0}, \tau_1^{w_1}, \dots, \tau_{M-1}^{w_{M-1}} \rangle)$$
$$= \tau'^{w'_j}_j \text{ such that: } \left| \sum_{p=0}^{j-1} w'_p - \sum_{p=j+1}^{M'-1} w'_p \right| < \varepsilon$$
$$= m(\tau_i^{w_i}, \mathcal{L}_{M-1}) \text{ with } w_i = 1$$
$$= m(\tau_i, \mathcal{L}_{M-1})$$

with $m(\tau_i, \mathcal{L}_{M-1})$ a GSM (or a composition of GSMs) applied to an element of the initial collection $\mathcal{L}_M$ and where $\sum_{p=0}^{j-1} w'_p$ ($\sum_{p=j+1}^{M'-1} w'_p$ respectively) is the sum $\mathscr{S}_1$ ($\mathscr{S}_2$ respectively) of the weights that are before (after respectively) element $\tau'^{w'_i}_j$.

$\varepsilon$ has to be negligible ($\varepsilon \ll w_i$), so both sums $\mathscr{S}_1$ and $\mathscr{S}_2$ must be as close as possible. $\varepsilon$ is a *negligible semantic gap, i.e.* a negligible difference between two linguistic descriptions of an object. The chosen method is to change scale, *i.e.* to subdivise the element into sub-elements. This way, a new collection is obtained and the sums $\mathscr{S}$ can be computed again. Thus the result of the aggregation is either an element of $\mathcal{L}_M$, or a *sub-element*. A sub-element is an element on which one or more GSMs have been applied.

### 2.2. Expressing the doubt linguistically

In [10], we focused on the various ways to express the hesitation or the doubt and we obtained two families of linguistic statements. The binary ones are: *between* $\tau_\alpha$ *and* $\tau_\beta$*;* $\tau_\alpha$ *or* $\tau_\beta$ (there is no condition on $\tau_\alpha$ nor $\tau_\beta$, *i.e.* they don't need to be subsequent values); the unary ones are: *at most* $\tau_\alpha$*; at least* $\tau_\alpha$*; everything except* $\tau_\alpha$.

In the statement "I hesitate between two alternatives but I tend to lean toward the second one", the words "tend to lean towards" add obviously a notion of a weight assigned on the second alternative (the second alternative is assigned a higher weight than the first one).

We have seen that the SWM permits to express an aggregation of a set of weighted alternatives. Considering $M$ alternatives denoted $\tau_0$ to $\tau_{M-1}$, the above statement representing an expert's opinion can be expressed with the following collection of weighted alternatives: $\langle \tau_0^{w_0}, \dots, \tau_i^{w_i}, \dots, \tau_{M-1}^{w_{M-1}} \rangle$. The weights $w_i$ permit to express the hesitation, with $\Sigma w_i = 1$.

The choice of the best weights to sum up the expert's opinion is considered below, according to the linguistic assessment.

So we assume in this paper that the SWM is able to express all kinds of hesitation, because they all are a question of weights on alternatives.

### 2.3. Hesitant fuzzy linguistic term sets

Recently, several scientists focused on hesitant linguistic expressions and formalisms dedicated to them. Torra introduced the concept of Hesitant Fuzzy Sets as extensions and generalizations of fuzzy sets [1]. Hesitant fuzzy sets deal with quantitative settings.

Besides, Hesitant Fuzzy Linguistic Term Sets, as *qualitatives settings*, have been proposed to provide a linguistic and computational basis to increase the richness of linguistic elicitation based on the fuzzy linguistic approach [2].

A *Hesitant Fuzzy Linguistic Term Set* (HFLTS), denoted $H_S$, is defined as an ordered finite subset of the consecutive linguistic terms of a linguistic term set.

Many basic operations have been defined such as upper bound and lower bound, the complement of an HFLTS, union and intersection between two HFLTS. The concept of *fuzzy envelope* of an HFLTS is also introduced as being a linguistic interval bounded by the minimal and the maximal elements of the HFLTS.

Several expressions are suggested to explicit the hesitation. These expressions are generated from a *context-free grammar* denoted $G_H$ with primary and composite terms, unary and binary relations and conjunction. Three expressions obtained through a transformation function $E_{G_H}$ are given: *at least, at most* and *between ...and ...* (see [16]).