IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 3, 2019, accepted December 13, 2019, date of publication December 25, 2019, date of current version January 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2962420

Entropy of Polysemantic Words for the Same Part of Speech

MIHAELA COLHON[1], FLORENTIN SMARANDACHE[2], AND DAN VALERIU VOINEA[3]
[1]Department of Computer Science, Faculty of Sciences, University of Craiova, 200585 Craiova, Romania
[2]Department of Mathematics, The University of New Mexico, Albuquerque, NM 87131, USA
[3]Faculty of Letters, University of Craiova, 200585 Craiova, Romania

Corresponding author: Mihaela Colhon (mcolhon@inf.ucv.ro)

ABSTRACT In this paper, a special type of polysemantic words, that is, words with multiple meanings for the same part of speech, are analyzed under the name of neutrosophic words. These words represent the most difficult cases for the disambiguation algorithms as they represent the most ambiguous natural language utterances. For approximate their meanings, we developed a semantic representation framework made by means of concepts from neutrosophic theory and entropy measure in which we incorporate sense related data. We show the advantages of the proposed framework in a sentiment classification task.

INDEX TERMS Neutrosophic sets, semantic word representation, sentiment classification.

I. INTRODUCTION

Every natural language word can have multiple realisations from the part-of-speech point of view, and for each of its possible parts-of-speech, it can have multiple meanings (especially the English words). Each sense creates a "sub-dimension" in the word's space determined by the part-of-speech to which it belongs in the given statement. The polysemantic words (words with multiple senses) can be described by several spaces (one space for each possible part-of-speech) and each space can include several subspaces determined by the meanings the word can have. In this manner, every dimension describes a certain facet of the analysed word. It is also true that certain senses are more frequent than others and in this manner they can force a certain facet to be more prominent than others.

We need a comprehensive and unitary study for natural language words formulated as a Multicriteria Decision Making problem [1] in which uncertainty is inevitably involved due to the subjectivity of humans [2]. It has been shown that different senses of the same word usually imply different sentiment orientations for the word under analysis. For instance, the word "good": in "good man" produces a positive utterance while in "good fight" indicates a negative statement. As a direct consequence we need studies that address both the interaction between word sense disambiguation and sentiment analysis. These are quite new studies in the literature as the researchers in this area must be intrigued by the usability of sense level information in sentiment analysis. Some researchers take this approach and compute the polarity score for each word sense [3], [4].

The present paper proposes a novel approach for word sentiment classification by extracting a set of semantic data from the SentiWordNet in order to compute a final estimation of the word polarity. SentiWordNet [3], [5] is a well-known freely available lexical resource for sentiment analysis which annotates each sense of a word with three polarity scores. These polarity scores represent the positivity, objectivity and negativity degrees of the annotated word sense ranging from 0 to 1 with their sum up to one. SentiWordNet (SWN) was built on the semantically-oriented WordNet [6], [7], which in its primary form, that is for English language, comprises 155287 words and 117659 senses.

There are two main approaches for sentiment analysis: machine learning and knowledge-based. From the machine learning perspective, the Support Vector Machines (SVM) classification (see, for example, [8], [9]) has the best classification performance for sentiment analysis [10], [11] outperforming both the Naïve Bayes and Maximum entropy classification methods. The knowledge-based methods usually make use of the most common sense of the words and in this manner an improvement of accuracy over the baseline was observed [12]. Also, the overall polarities of different senses in each part-of-speech tag categories are also

The associate editor coordinating the review of this manuscript and approving it for publication was Pasquale De Meo.

determined [13]. However, the commonly used n-gram features are not robust enough and show widely varying behaviour across different domains [14].

The method we propose in this paper offers a knowledge-based solution for semantic word representation which targets sentiment classification and makes use of the general concepts of neutrosophic theory and entropy measure. A previous study that applies neutrosophic theory in sentiment analysis domain is given in [15]. In this paper we concentrate our approach by keeping in mind only the most difficult cases for sentiment classification. They are represented by a special class of polysemantic words with different meanings for the same part-of-speech realisation. In the present paper these words are named *neutrosophic words* because their representation involves the core concepts of neutrosophic theory.

With this article we are in line with the neutrosphic word representations firstly proposed in [16] and then refined in [17] in which the SentiWordNet (shortly SWN) sentiment scores are interpreted as truth-fullness degrees. The study proposed in this paper also makes use of the SWN polarity scores of each word's sense, this time in order to determine the overall sentiment score value. The involved computations apply entropy on the words' sentiment scores as a measure of disorder for the words' polarities.

The paper is organised as follows: the Related Works section overviews the most commonly used multi-space representation techniques in neutrosophy. Section III presents the proposed semantic-level representation which treats the words as union of neutrosophic sets. In Section IV we show how this type of representation can be used in conjunction with a sentiment analysis study. Section V exemplifies all the involved theoretical concepts on a study case also providing the obtained results and the last section is dedicated to the conclusions and our future directions.

II. RELATED WORKS

The concept of multi-space was introduced by Smarandache in 1969 [18] by following the idea of hybrid mathematics - especially hybrid geometry [19], [20] for combining different fields into a unifying field [21]–[24].

Let Ω be a universe of discourse and a subset $S \subseteq \Omega$. Let $[0, 1]$ be a closed interval and three subsets $T, I, F \subseteq [0, 1]$. Then, a relationship of an element $x \in S$ with respect to the subset S is $x(T, I, F)$, which means the following: the *confidence set* of x is T, the *indefinite set* of x is I, and the *failing set* of x is F. A set S, together with the corresponding three subsets T, I, F for each element x in S, is said to be a *neutrosophic set* [19], [25].

Let Σ be a set and $A_1, A_2, \ldots, A_k \subseteq \Sigma$. Define $3k$ functions $f_1^z, f_2^z, \ldots, f_k^z$ by $f_i^z : A_i \to [0, 1]$, $1 \le i \le k$, where $z \in \{T, I, F\}$. If we denote by $(A_i; f_i^T, f_i^I, f_i^F)$ the subset A_i together with three functions f_i^T, f_i^I, f_i^F, $1 \le i \le k$, then [19]:

$$\bigcup_{i=1}^{k}(A_i; f_i^T, f_i^I, f_i^F)$$

is a union of neutrosophic sets which are generalizations of classical sets.

Indeed, if we take $f_i^T = 1, f_i^I = f_i^F = 0$ for $i = \overline{1, k}$ we obtain [19]:

$$\bigcup_{i=1}^{k}(A_i; f_i^T, f_i^I, f_i^F) = \bigcup_{i=1}^{k} A_i$$

and correspondingly, for $f_i^T = f_i^I = 0, f_i^F = 1$, $i = \overline{1, k}$ we obtain the complementary sets [19]:

$$\bigcup_{i=1}^{k}(A_i; f_i^T, f_i^I, f_i^F) = \overline{\bigcup_{i=1}^{k} A_i}$$

The appurtenance and non-appurtenance is obtained if there is an integer s such that $f_i^T = 1, f_i^I = f_i^F = 0, 1 \le i \le s$, but $f_j^T = f_j^I = 0, f_j^F = 1, s + 1 \le j \le k$.

$$\bigcup_{i=1}^{k}(A_i; f_i^T, f_i^I, f_i^F) = \bigcup_{i=1}^{s} A_i \cup \overline{\bigcup_{i=s+1}^{k} A_i}$$

The general neutrosophic set is obtained if there is an integer l such that $f_l^T \neq 1$ for $1 \le l \le s$, or $f_l^F \neq 1$ for $s < l \le n$. The resulted union cannot be represented by abstract sets.

III. SEMANTIC-LEVEL REPRESENTATION FOR WORDS

As we have already pointed out in the Introduction section, a word is not a simple data, it can have several (syntactic) attributes and can support more than one semantic interpretations. Metaphorically speaking a word is like a diamond: it can brighter a life or, by contrary, it can cut and destroy. But, from our study point of view, a word is just an entity that can have multiple semantic facets.

As we have already pointed out, a word can have more than one part-of-speech, like the word ''good'' which can be adjective, noun or adverb and to which we dedicate an extensive study in the Section V. There are programs that can automatically identify the part-of-speech of a certain word in a given context. These programs are named Part-Of-Speech Taggers and for most of the languages their accuracy is quite high (more than 90%).

On contrary, determining the meaning of a polysemous word in a specific context - that is, performing a disambiguation on the word's senses, can be a laborious task. In spite of the great number of existing disambiguation algorithms, the problem of word sense disambiguation remains an open one [26]. For some languages like English the accuracy of the disambiguation algorithms does not overcome 75%.

It is obviously that we need to model indeterminacy in the semantic word representations. This is the reason why, in the present study we choose to model word representations using neutrosophic theory as, in contrast to intuitionistic fuzzy sets and also interval valued intuitionistic fuzzy sets, indeterminacy degree of an element is explicitly expressed by the neutrosophic sets [27]. Moreover, in [29] the authors state that single valued neutrosophic (SVN) set is a better tool to deal with incomplete, inconsistent and indeterminate information than fuzzy set (FS) and intuitionistic FS (IFS). With the present study we are in line with these assumptions continuing also our previous works in which the natural language words are modelled as single-valued neutrosophic sets in order to approximate their ambiguous meaning [16], [17].

In the representation we propose in this paper a word can have multiple dimensions organised on several plans:

- the POS plans are determined by the possible part of speech data of the word
- each POS plan can have several sense units, determined by the possible word's senses under that POS data
- finally, each sense unit is made of some components (sentiment scores) which describe the sense meaning polarity

A. WORDS AS UNION OF NEUTROSOPHIC SETS

The first step in creating a semantic representation is to decide what features to use and how to encode these features. From the features set a word can have, in this study we consider the part-of-speech as the syntactic feature and the word's sense(s) as its semantic interpretation(s).

In what follows, we name semantic facets or simply facets - all the word's data based on which the semantic interpretation can be defined. Using concepts from neutrosophic sets theory [30] we propose the following semantic representation of a word.

Definition 1: The semantic representation of a word by means neutrosophic theory concepts is defined as:

$$w = \bigcup_{i=1}^{k} (sense_i; f_i^T, f_i^I, f_i^F)$$

where:

- k represents the number of senses the word can have
- $f_i^T, f_i^I f_i^F \;:\; Facets \to [0, 1]$ are the membership functions for the $sense_i$, $i = \overline{1, k}$, such that:
  - f_i^T represents the membership degree,
  - f_i^I represents the indeterminacy degree and
  - f_i^F is the degree of nonmembership degree
- *Facets* set includes all the data that characterise the word from the semantic point of view.

In this assertion, a word becomes a union of neutrosophic sets. For the ith sense of the word w, the membership functions of the word's semantic facets fulfil the following properties:

$$\forall x \in Facets : f_i^T(x) + f_i^I(x) + f_i^F(x) = 1 \qquad (1)$$

and if $Facets = \{x_1, \ldots, x_m\}$ then:

$$\sum_{j=1}^{m} f_i^T(x_j) + f_i^I(x_j) + f_i^F(x_j) = m \qquad (2)$$

In order to include the information about the part-of-speech data (shortly POS data) we need to refine the representation given in Definition 1. We consider the general case in which a word can have n possible parts of speech $POS_1, \ldots, POS_n$, with $n \geq 1$, and for each part of speech POS_j the word can have k_j senses, $k_j \geq 1$. The representation given in Definition 1 becomes:

$$w = \bigcup_{i=1}^{k_1} (sense_{i;POS_1}; f_{i;POS_1}^T, f_{i;POS_1}^I, f_{i;POS_1}^F) \cup \ldots$$
$$\ldots \cup \bigcup_{i=1}^{k_n} (sense_{i;POS_n}; f_{i;POS_n}^T, f_{i;POS_n}^I, f_{i;POS_n}^F) \qquad (3)$$

Using the representation given in Equation 3, the senses corresponding to a certain part of speech POS_j with $j \in \{1, \ldots n\}$, can be obtained as follows:

$$\begin{aligned}(w)_{POS_j} &= w \cap (w)_{POS_j} \\ &= \bigcup_{i=1}^{k_j} (sense_{i;POS_j}; f_{i;POS_j}^T, f_{i;POS_j}^I, f_{i;POS_j}^F) \qquad (4)\end{aligned}$$

Furthermore, we can apply another filtering on word representation given in Equation 4 if we consider the case in which a specific sense of the word w results to be realised in a given context. Let us note this sense with $sense_{m;POS_j}$ with $m \in \{1, k_j\}$. By applying concepts from neutrosophic sets theory we obtain:

$$f_{m;POS_j}^T = 1,\; f_{m;POS_j}^I = f_{m;POS_j}^F = 0 \quad \text{and}\; f_{l;POS_j}^T = f_{l;POS_j}^I = 0,$$
$$f_{l;POS_j}^F = 1 \quad \text{for}\; l \neq m, \quad l, m = \overline{1, k_j}$$

which implies:

$$\begin{aligned}(w)_{POS_j} &= \bigcup_{i=1}^{k_j} (sense_{i;POS_j}; f_{i;POS_j}^T, f_{i;POS_j}^I, f_{i;POS_j}^F) \\ &= (sense_{m;POS_j}; 1, 0, 0) \cup \bigcup_{l \neq m} (sense_{l;POS_j}; 0, 0, 1) \\ &= sense_{m;POS_j} \cup \overline{\bigcup_{l \neq m} sense_{l;POS_j}} \\ &= sense_{m;POS_j} = (m\text{-}th\; sense\; of\; w)_{POS_j} \qquad (5)\end{aligned}$$

The representation given in Equation 5 corresponds to the most unambiguous case, more precisely to the situation in which we know both the word's part of speech (noted here with POS_j) and the word sense (noted with $sense_{m;POS_j}$).

But, the problems with natural language processing comes from ambiguity - when we could not identify (using automatic tools) which sense is realised in the given context from the set of the word's possible senses (noted here with $\cup_{i=1}^{k_j} sense_{i;POS_j}$). This ambiguity case is depicted by the general case given in Equation 3.

In what follows we will use a simplified form of Equation 3 in which POS_j data is removed from the annotations sequences corresponding to the senses and membership functions. Thus, Equation 3 becomes:

$$(w)_{POS_j} = \bigcup_{i=1}^{k_j} (sense_i; f_i^T, f_i^I, f_i^F) \qquad (6)$$

In the next section we present a method by means of which we can eliminate the "noises" from an ambiguous semantic word representation, more precisely, a representation that includes more than one possible sense. We resolve these issues using Neutrosophic Theory and Entropy measure. Our proposal is described in conjunction with a sentiment analysis study in which the semantic word representation has the form of a three sentiment scores tuple.

IV. WORD SEMANTIC REPRESENTATION WITH SENTIMENT SCORES

Sense discrimination addresses words with multiple senses and is done in conjunction with a particular context in which only one sense is realised. This analysis has a semantic nature and is quite difficult to perform it using automatic tools, especially if the realisation context is poor in information that could filter the correct word meaning from the set of possible ones. In order to exemplify our proposal we choose to interpret the word semantics from a sentiment analysis the point of view. Thus, each sense of a word will be represented using its sentiment scores.

In what follows, let us consider the approach firstly proposed in [16] and then extended in [17] in which a word w is interpreted as a single-value neutrosophic set constructed upon its sentiment scores which describe the word's sense-level polarity information being denoted in what follows with (sc_+, sc_0, sc_-), where:

- sc_+ denotes the word positive score,
- sc_0 represents the word neutral score and
- sc_- stands for the word negative score.

As in [16] and [17] we use SentiWordNet lexical resource for providing the required information for the sentiment scores of the English words.

For a word w with k_j senses under a POS_j part-of-speech realisation, the semantic representation is defined as a union of the tuples: $sense_i = (sc_{+_i}, sc_{0_i}, sc_{-_i})$ with $i \in \{1, \ldots, k_j\}$. The Equation 6 becomes:

$$(w)_{POS_j} = \bigcup_{i=1}^{k_j} ((sc_{+_i}, sc_{0_i}, sc_{-_i}); f_i^T, f_i^I, f_i^F) \tag{7}$$

with $sc_{+_i}, sc_{0_i}, sc_{-_i} \in [0, 1]$. The semantic representation given in Equation 7 implies that each word's sense will include three facets: the positive, the neutral and the negative one. By preserving the notation where $+$ stands for positive, 0 for neutral and $-$ for negative facet, we take $Facets = \{+, 0, -\}$.

The representation given in Equation 7 can be rewritten as:

$$(w)_{POS_j} = \bigcup_{i=1}^{k_j} ((sc_{+_i}, sc_{0_i}, sc_{-_i}); (\{f_i^T(x)\}_{x\in Facets}), (\{f_i^I(x)\}_{x\in Facets}), (\{f_i^F(x)\}_{x\in Facets})) \tag{8}$$

where $f_i^T(x)$, $f_i^I(x)$ and $f_i^F(x)$ represents the membership functions corresponding to the facet x of the ith sense, $x \in Facets$ and $(\{f_i^M(x)\}_{x\in Facets})$ briefly notes the membership functions $\begin{pmatrix} f_i^M(+) \\ f_i^M(0) \\ f_i^M(-) \end{pmatrix}$, $M \in \{T, I, F\}$.

Remark: For the representation given in Equation 8, the default case corresponds to the maximum certainty case where no imprecision occurs which, in terms of membership function is depicted by $f_i^T(\{+ \mid 0 \mid -\}_i) = 1, f_i^I(\{+ \mid 0 \mid -\}_i) = 0, f_i^F(\{+ \mid 0 \mid -\}_i) = 0, i = \overline{1, k_j}$.

We preface the study that addresses the multi-facets words by enumerating the form in which the *one facet words* are represented in our proposal. These words are the extreme cases of our study and every neutrosophic study provides them.

Case 1: If $sc_{+_i} = 1$, $sc_{0_i} = sc_{-_i} = 0$ and $f_i^T(\{+ \mid 0 \mid -\}_i) = 1, f_i^I(\{+ \mid 0 \mid -\}_i) = 0, f_i^F(\{+ \mid 0 \mid -\}_i) = 0$ for every $i = \overline{1, k_j}$ then:

$$(w)_{POS_j} = \bigcup_{i=1}^{k_j} \left((1, 0, 0); \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) = (1, 0, 0)$$

The interpretation of Case 1 is: for all the senses corresponding to the POS_j part-of-speech the word w is *pure positive*.

Case 2: If $sc_{+_i} = sc_{0_i} = 0$, $sc_{-_i} = 1$ and $f_i^T(\{+ \mid 0 \mid -\}_i) = 1, f_i^I(\{+ \mid 0 \mid -\}_i) = 0, f_i^F(\{+ \mid 0 \mid -\}_i) = 0$ for every $i = \overline{1, k_j}$ then:

$$(w)_{POS_j} = \bigcup_{i=1}^{k_j} \left((0, 0, 1); \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) = (0, 0, 1)$$

The interpretation of Case 2 is: for all the senses corresponding to the POS_j part-of-speech the word w is *pure negative*.

Case 3: If $sc_{+_i} = sc_{-_i} = 0$, $sc_{0_i} = 1$ and $f_i^T(\{+ \mid 0 \mid -\}_i) = 1, f_i^I(\{+ \mid 0 \mid -\}_i) = 0, f_i^F(\{+ \mid 0 \mid -\}_i) = 0$ for every $i = \overline{1, k_j}$ then:

$$(w)_{POS_j} = \bigcup_{i=1}^{k_j} \left((0, 1, 0); \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) = (0, 1, 0)$$

The interpretation of Case 3 is: for all the senses corresponding to the POS_j part-of-speech the word w is *pure neutral*.

These three cases correspond to the non-ambiguous words, that is, words with a unique sense (one semantic representation) or similar semantic representations for all of their possible senses.

Since in a natural language there are many words (especially in English) with multiple senses - the *polysemantic words*, in what follows we will concentrate our study only on these words. For the polysemantic words we get different semantic representations that must be resolved by dealing with many degrees of uncertainties. In this case, the simple reunion of their semantic dimensions is a general neutrosophic set that cannot be formalised using abstract set theories. For this reason, in the next definition we introduce the concept of *neutrosophic word* in conjunction with a sentiment analysis.

Definition 2: A *neutrosophic word* is a polysemantic word that under the same part of speech realization has at least two different sentiment polarities which means:

$(\exists(w)_{POS_j}$ with $k_j > 1$ senses$) \wedge (\exists i_1, i_2 \in \{1, \ldots, k_j\}, i_1 \neq i_2$: $sense_{i_1} \neq sense_{i_2})$

Different sense tuples imply different sentiment scores and we obtain:

$(\exists(w)_{POS_j}$ with $k_j > 1$ senses$) \wedge [\exists i_1, i_2 \in \{1, \ldots, k_j\}, i_1 \neq i_2$: $(sc_{+_{i_1}}, sc_{0_{i_1}}, sc_{-_{i_1}}) \neq (sc_{+_{i_2}}, sc_{0_{i_2}}, sc_{-_{i_2}})]$

As a direct consequence, the semantic representation of neutrosophic words is:

$$(w)_{POS_j} = \bigcup_{i \in \{i_1, i_2, \ldots\}} ((sc_{+_i}, sc_{0_i}, sc_{-_i}); (\{f_i^T(x)\}_{x \in Facets}), (\{f_i^I(x)\}_{x \in Facets}), (\{f_i^F(x)\}_{x \in Facets}))$$

with $sc_{+_{i_1}} \neq sc_{+_{i_2}}$ or $sc_{0_{i_1}} \neq sc_{0_{i_2}}$ or $sc_{-_{i_1}} \neq sc_{-_{i_2}}$ and $f_{i_1}^T(\{+ \mid 0 \mid -\}), f_{i_2}^T(\{+ \mid 0 \mid -\}) > 0$, $i_1 \neq i_2$. By the fact that the membership degrees are greater than 0, we obtain for a neutrosophic word *w the necessity of having (at least) two different sentiment representations* for the same $(w)_{POS_j}$.

The neutrosophic theory means from the very beginning dealing with uncertainty. This is also true for the neutrosophic words. These words can be evidenced in case of an imprecise disambiguation mechanism which fails in recognising what sense is realised in the given context even if the part-of-speech data is correctly provided.

In our approach, a neutrosophic word is synonym with a word that has different sense facets and for which we cannot establish a unique semantic representation. For the chosen sentiment analysis exemplification, different sense facets for a word means different sentiment scores tuples.

In the next section we exemplify how the proposed method works. We show that using the neutrosophic sets theory and applying the entropy measure on the word representations we can identify the word's sentiment facet with respect to the given part-of-speech.

A. ENTROPY AS A MEASURE OF UNCERTAINTY FOR THE NEUTROSOPHIC WORDS REPRESENTATIONS

Fuzzy entropy, distance measure and similarity measure are three basic concepts used in fuzzy sets theory [27]. Among them, Entropy is an efficient tool to model uncertainty [28] or, in layman terms, Entropy is a measure of disorder. It can be used in order to measure how disorganised an input values set is by calculating the entropy of their values/labels. Entropy is high if the input values are highly varied and low if many input data have the same value. In mathematical terms, Entropy is defined as the sum of the probability of each input values or labels times the log probability of that label:

$$E(labels) = -\sum_{l \in labels} P(l) log_2 P(l) \tag{9}$$

where $P(l)$ is the frequency probability of the *label* item in the considered data and *labels* denotes the set of possible labels.

From this definition we obtain that labels with low frequency do not affect much the entropy (because $P(l)$ is small). The same result for labels with high frequency as in their case, $log_2 P(l)$ is small. Only when the inputs have wide varieties of labels (and as a direct consequence, these many labels have a medium frequency) the entropy is high because neither $P(l)$ nor $log_2 P(l)$ is small.

Entropy has values between 0 and 1 and high entropy values stand for high levels of disorder or ''low level of purity''. Following this property, we can qualify the uncertainty of the words' semantic nature by applying the entropy measure on their sense representation labels: *the higher the values for entropy measure the higher the level of uncertainty for the analyzed word representations*.

The neutrosophic word is a concept with more than one possible sense for at least one of its possible part-of-speech data. On the other hand, entropy is a measure of uncertainty. Between the possible senses we can have certain similarity degrees and the entropy measure can be used in order to determine how similar or dissimilar these senses are.

The most common manner to unify a set of possible representations into a single one is to consider only the maximum (or the minimum) value or to average the values (in our case, the sentiment scores) as in the following formula:

$$Avg\left(\bigcup_{i=1}^{k_j}(sc_{+_i}, sc_{0_i}, sc_{-_i})\right) = \left(\frac{1}{k_j}\sum_{i=1}^{k_j} sc_{+_i}, \frac{1}{k_j}\sum_{i=1}^{k_j} sc_{0_i}, \frac{1}{k_j}\sum_{i=1}^{k_j} sc_{-_i}\right) \tag{10}$$

where k_j notes the number of senses for the analysed word. But this method of unifying different representation can be trustful only if the averaged values are not very dissimilar with the initial ones.

Example 1: Let us consider a word w with two extreme sentiment scores tuples: $(0, 0, 1)$ and $(1, 0, 0)$. Overall, we obtain two different facets: in the first representation we have a *pure positive word* while in the second we get a *pure negative word*. If we merge these two representation by averaging their sentiment scores values we get $(0.5, 0, 0.5)$ - a representation that could be interpreted as a *neutral word*. Definitely this would be a wrong classification for a strong sentiment word.

We define a bijective mapping for labelling the sentiment score values to a set of three strength degrees, $SD = \{low, medium, high\}$. We obtain $sd : [0, 1] \rightarrow SD$ with:

$$sd(score) = \begin{cases} low, & \text{if } score < 0.4 \\ medium, & \text{if } score \in [0.4, 0.6] \\ high, & \text{if } score > 0.6 \end{cases}$$

Mapping the score values to the *SD* labels we get ''low'' label for a small score, ''medium'' for not a small but also not a high score and ''high'' for a big score. Using these strength degrees we can qualify by means of the entropy measure calculated as in Equation 9 how disorganised the scores values are from the point of view of the sentiment strength. All the involved operations are given in Algorithm 1.

Algorithm 1 Merging Multiple Semantic Representations of a Neutrosophic Word $(w)_{POS_j}$

```
INPUT: ∪_{i=1}^{k_j}(sc_{+_i}, sc_{0_i}, sc_{-_i})
for each x in Facets:
    Entropy(x) ← E(∪_{i=1}^{k_j} sd(sc_{x_i}))
    Avg(x) ← Avg(∪_{i=1}^{k_j} sc_{x_i}) ← 1/k_j ∑_{i=1}^{k_j} sc_{x_i}
    f^T(x) ← 1 − Entropy(x)
    f^I(x) ← Entropy(x)
    f^F(x) ← 0
endfor
OUTPUT: ∪_{x∈Facets} Avg(x), f^T(x), f^I(x), f^F(x)
```

We can now give the manner in which the multiple representations of a neutrosophic word $(w)_{POS_j}$ can be unified into a unique sentiment representation $Avg(w)_{POS_j}$ based on the values provided by Algorithm 1:

$$\begin{aligned}
&Avg((w)_{POS_j})\\
&= Avg\left(\bigcup_{i=1}^{k_j}\left((sc_{+_i}, sc_{0_i}, sc_{-_i}); \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right)\right)\\
&= (\Big(Avg(\cup_{i=1}^{k_j} sc_{+_i}), Avg(\cup_{i=1}^{k_j} sc_{0_i}), Avg(\cup_{i=1}^{k_j} sc_{-_i})\Big);\\
&\quad (\{f^T(x)\}_{x\in Facets}), (\{f^I(x)\}_{x\in Facets}), (\{f^F(x)\}_{x\in Facets}))
\end{aligned} \tag{11}$$

In Algorithm 1 we model the degrees of trustfulness for the resulted average scores representation by means of the membership functions, such that $\forall x \in Facets$:

- If the entropy $Entropy(x)$ is small (the minimum value is 0) then the average value $Avg(x)$ can approximate with high degree of certainty the initial word's sentiment scores; in this case the membership function for the facet x is set to a big value (almost 1) as $f^T(x) \leftarrow 1 - Entropy(x)$.
- If the entropy $Entropy(x)$ is high (the maximum value is 1) then the membership function is set to a small value (almost 0) while the indeterminacy degree $f^I(x)$ is set to be equal with the entropy function value.
- For preserving the sum of the membership functions to value 1 (see Equation 1), the nonmembership degree $f^F(x)$ for the facet x is always 0.

For the case given in **Example 1** we obtain that the entropy corresponding to the positive and negative scores is equal to its maximum value: $E(+) = E(-) = 1$, while the entropy for the neutral scores is zero. The resulted average representation can be written as follows:

$$\begin{aligned}
Avg(w) &= ((Avg(\cup_{i=1}^{2} sc_{+_i}), Avg(\cup_{i=1}^{2} sc_{0_i}),\\
&\qquad Avg(\cup_{i=1}^{2} sc_{-_i})); f^T, f^I, f^F)\\
&= ((0.5, 0, 0.5); f^T, f^I, f^F)\\
&= \left((0.5, 0, 0.5); \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right)
\end{aligned} \tag{12}$$

The representation given in Equation 12 tells more about *what the word is not* than about the type the word *is* as we consider **Example 1** only for showing why the simple unification of multiple representations by averaging their values is not always enough. As one can observe, the representation given in Equation 12 tells with maximum certainty that the word is not a neutral word. For the obtained positive and negative scores the indeterminacy membership functions have maximum values, illustrating in this way a maximum indeterminacy degree. This extreme case is quite rarely, being presented only for its theoretical purpose.

In the next section we apply the proposed method on a real data: a neutrosophic word in its all possible parts of speech. With this complex case we show that the method described in this article succeeds in merging multiple and diverse semantic word representations.

V. STUDY CASE

The word "good" appears in WordNet with three different parts of speech (noun, adjective, and adverb) and with many senses for each of its syntactic labels. We consider this word represents a perfect example for the neutrosophic word concept introduced in this paper and for this reason we dedicate the study case to it.

In Table 1 are given all the senses the word "good" can have, grouped upon the part-of-speech data. Each sense is given together with the sentiment scores extracted from SentiWordNet and also with its definition and some examples (as they are given in WordNet).

In Table 2 we gather all the data extracted from SentiWordNet: the word's parts of speech, the three facets given by the corresponding sentiment scores and the distributions among the senses of the sentiment scores. We also give the entropy measures for each word's facet in all the three parts of speech and also the average values of the sentiment scores.

By applying Algorithm 1 on the SentiWordNet scores of the word "good" we obtain the following representations (see also Table 2):

$$\begin{aligned}
&Avg((good)_{ADJ})\\
&\quad = (\Big(Avg(\cup_{i=1}^{21} sc_{+_i}), Avg(\cup_{i=1}^{21} sc_{0_i}), Avg(\cup_{i=1}^{21} sc_{-_i})\Big);\\
&f^T_{ADJ}, f^I_{ADJ}, f^F_{ADJ})\\
&\quad = \Big((0.61, 0.38, 0); f^T_{ADJ}, f^I_{ADJ}, f^F_{ADJ}\Big)\\
&\quad = \left((0.61, 0.38, 0); \begin{pmatrix}0.59\\0.59\\1\end{pmatrix}, \begin{pmatrix}0.41\\0.41\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right)
\end{aligned} \tag{13}$$

$$\begin{aligned}
&Avg((good)_{NOUN})\\
&\quad = (\Big(Avg(\cup_{i=1}^{4} sc_{+_i}), Avg(\cup_{i=1}^{4} sc_{0_i}), Avg(\cup_{i=1}^{4} sc_{-_i})\Big);\\
&\quad f^T_{NOUN}, f^I_{NOUN}, f^F_{NOUN})\\
&\quad = \Big((0.5, 0.5, 0); f^T_{ADJ}, f^I_{ADJ}, f^F_{ADJ}\Big)\\
&\quad = \left((0.5, 0.5, 0); \begin{pmatrix}0.25\\0.25\\1\end{pmatrix}, \begin{pmatrix}0.75\\0.75\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right)
\end{aligned} \tag{14}$$

TABLE 1. The SentiWordNet data for the word "good".

POS	Sentiment Scores	Example
Noun	(0.5, 0.5, 0)	benefit; "for your own good"; "what's the good of worrying?"
	(0.875, 0.125, 0)	moral excellence or admirableness; "there is much good to be found in people"
	(0.625, 0.375, 0)	that which is pleasing or valuable or useful; "weigh the good against the bad"; "among the highest goods of all are happiness and self-realization"
	(0, 1, 0)	articles of commerce
ADV	(0.375, 0.625, 0)	in a good or proper or satisfactory manner or to a high standard; "the baby can walk pretty good"
	(0, 1, 0)	completely and absolutely; "he was soundly defeated"; "we beat him good"
ADJ	(0.75, 0.25, 0)	having desirable or positive qualities especially those suitable for a thing specified; "good news from the hospital"; "a good report card"
	(0, 1, 0)	having the normally expected amount; "gives full measure"; "gives good measure"
	(1, 0, 0)	morally admirable
	(1, 0, 0)	deserving of esteem and respect; "ruined the family's good name"
	(0.625, 0.375, 0)	promoting or enhancing well-being; "the experience was good for her"
	(1, 0, 0)	agreeable or pleasing; "we all had a good time"; "good manners"
	(0.75, 0.25, 0)	of moral excellence; "a genuinely good person"
	(0.625, 0.375, 0)	having or showing knowledge and skill and aptitude; "a good mechanic"
	(0.625, 0.375, 0)	thorough; "had a good workout"; "gave the house a good cleaning"
	(0.5, 0.5, 0)	with or in a close or intimate relationship; "a good friend"
	(0.5, 0.5, 0)	financially sound; "a good investment"
	(0.375, 0.625, 0)	most suitable or right for a particular purpose; "a good time to plant tomatoes"
	(0.625, 0.375, 0)	resulting favorably; "it's a good thing that I wasn't there"; "it is good that you stayed"
	(0, 1, 0)	exerting force or influence; "a warranty good for two years"
	(0.625, 0.375, 0)	capable of pleasing; "good looks"
	(0.75, 0.25, 0)	appealing to the mind; "good music"
	(0.75, 0.25, 0)	in excellent physical condition; "good teeth"; "I still have one good leg"
	(0.875, 0.125, 0)	tending to promote physical well-being; beneficial to health; "a good night's sleep"
	(0.5, 0.5, 0)	not forged; "a good dollar bill"
	(0.375, 0.5, 0.125)	not left to spoil; "the meat is still good"
	(0.75, 0.25, 0)	generally admired; "good taste"

$$
\begin{aligned}
&Avg((good)_{ADV})\\
&\quad = \left(\left(Avg(\cup_{i=1}^{2} sc_{+_i}), Avg(\cup_{i=1}^{2} sc_{0_i}), Avg(\cup_{i=1}^{2} sc_{-_i})\right);\right.\\
&\qquad \left. f_{ADV}^{T}, f_{ADV}^{I}, f_{ADV}^{F}\right)\\
&\quad = \left((0.18, 0.81, 0); f_{ADV}^{T}, f_{ADV}^{I}, f_{ADV}^{F}\right)\\
&\quad = \left((0.18, 0.81, 0); \begin{pmatrix}0.59\\0.59\\1\end{pmatrix}, \begin{pmatrix}0.41\\0.41\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right)
\end{aligned} \tag{15}
$$

These results can be interpreted as follows: no matter its part of speech realisation, we can precisely say that the word "good" is NOT a negative word. Two possible facets remain: the positive and the neutral. From the results obtained in Equations 13 and 15 we can conclude:

TABLE 2. The semantic representations of the word "good". The negative scores, being not representative (the greatest value is 0.12), are omitted in the listing.

	Adjective			Noun			Adverb		
	sc_+	sc_0	#senses	sc_+	sc_0	#senses	sc_+	sc_0	#senses
Sent. Scores	0.75	0.25	5	0.5	0.5	1	0.37	0.62	1
	0	1	2	0.87	0.12	1	0	1	1
	1	0	3	0.62	0.37	1			
	0.62	0.37	5	0	1	1			
	0.5	0.5	3						
	0.37	0.62	1						
	0.87	0.12	1						
	0.37	0.5	1						
Entropy	0.41	0.41		0.75	0.75		0	0	
Avg Scores	0.61	0.38		0.5	0.5		0.18	0.81	

- the word "good" as adverb is a neutral word because its neutral average score is 0.81 with $f_{ADV}^{T}(0) = 0.59$, a value that exceeds by far its positive average score (0.18 with $f_{ADV}^{T}(+) = 0.59$)
- the word "good" as adjective is a positive word because its positive average score is 0.61 with $f_{ADJ}^{T}(+) = 0.59$ while the neutral average score is only 0.38, with $f_{ADJ}^{T}(0) = 0.59$

As a noun, we can consider it positive or neutral word, in both cases with high indeterminate degrees: $f_{NOUN}^{I}(+) = f_{NOUN}^{I}(0) = 0.75$, its average positive and neutral scores equal with 0.5 (see Equation 14). This is the case when additional filters taken from the context in which the word occurs must be applied in order to establish the word semantic facet.

VI. CONCLUSION AND FUTURE WORK

As pointed out in [31] each object has a corresponding (fuzzy, intuitionistic fuzzy, or neutrosophic) degree of appurtenance to a set of classification classes, with respect to its attributes' values.

In the present paper we propose a method that determines the appurtenance degrees of the semantic facets of a natural language word based on the entropy measure. We apply the proposed method on a real data: a polysemantic word in its all possible parts of speech. We prove with this complex study case that the method succeeds in merging multiple and diverse semantic word representations by filtering the "noises" through the entropy function values. The proposed method can be improved in case of high entropy values when additional filters must be applied by taken into account the word contextual data. The developing of these additional filters represents the trigger of our future studies.

REFERENCES

[1] F. Xiao, "A multiple–criteria decision–making method based on D numbers and belief entropy," *Int. J. Fuzzy Syst.*, vol. 21, no. 4, pp. 1144–1153, Jun. 2019.

[2] F. Xiao, "EFMCDM: Evidential fuzzy multicriteria decision making based on belief entropy," *IEEE Trans. Fuzzy Syst.*, to be published.

[3] A. Esuli and F. Sebastiani, "SentiWordNet: A publicly available lexical resource for opinion mining," in *Proc. LREC*, Genoa, Italy, 2006, pp. 417–422.
[4] J. Wiebe and R. Mihalcea, "Word sense and subjectivity," in *Proc. COLING/ACL*, Sydney, Australia, 2006, pp. 1065–1072.
[5] S. Baccianella, A. Esuli, and F. Sebastiani, "SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," in *Proc. LREC*, Valletta, Malta, 2010, pp. 2200–2204.
[6] G. A. Miller, "WordNet: A lexical database for English," *Commun. ACM*, vol. 38, no. 11, pp. 39–41, 1995.
[7] C. Fellbaum, *WordNet: An Electronic Lexical Database* Cambridge, MA, USA: MIT Press, 1998.
[8] S. Mohammad, S. Kiritchenko, and X. Zhu, "Nrc-canada: Building the state-of-the-art in sentiment analysis of tweets," in *Proc. SemEval*, Atlanta, GA, USA, 2013, pp. 321–327.
[9] M. C. Mihăescu, "Classification of users by using support vector machines," in*Proc. WIM*, Craiova, Romania, 2012, Art. no. 68.
[10] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Ret.*, vol. 2, nos. 1–2, pp. 1–135, 2008.
[11] K. Dave, S. Lawrence, and D. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *Proc. WWW*, Budapest, Hungary, 2003, pp. 519–528.
[12] E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi, "Knowledge–based approaches to concept–level sentiment analysis," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 12–14, Mar. 2013.
[13] C. Sumanth and D. Inkpen, "How much does word sense disambiguation help in sentiment analysis of micropost data?" in *Proc. 6th Workshop Comput. Approaches Subjectivity, Sentiment Social Media Anal.*, Berlin, Germany, 2015, pp. 115–121.
[14] Z. Fei, J. Liu, and G. Wu, "Sentiment classification using phrase patterns," in *Proc. IEEE Int. Conf. Comput. Inf. Technol.* Washington, DC, USA, Sep. 2004, pp. 1147–1152.
[15] F. Smarandache, M. Teodorescu, and D. Gîfu, "Neutrosophy, a sentiment analysis model," in *Proc. RUMOUR*, Toronto, ON, Canada, 2017, pp. 38–41.
[16] M. Colhon, Ş. Vlăduţescu, and X. Negrea, "How objective a neutral word is? A neutrosophic approach for the objectivity degrees of neutral words," *Symmetry*, vol. 9, no. 11, p. 280, Nov. 2017.
[17] F. Smarandache, M. Colhon, Ş. Vlăduţescu, and X. Negrea, "Word-level neutrosophic sentiment similarity," *Appl. Soft Comput.*, vol. 80, pp. 167–176, Jul. 2019.
[18] F. Smarandache, "Neutrosophic transdisciplinarity—Multi-space & multi-structure," in *Proc. State Archives*, Valcea Branch, Romania, 1969.
[19] L. Mao, *Smarandache Geometries & Map Theory with Applications (I)*, (English-Chinese Bilingual Edition). Beijing, China: Chinese Academy of Sciences, 2006.
[20] L. Mao, *Automorphism Groups of Maps, Surfaces and Smarandache Geometries*. Beijing, China: Beijing Institute of Civil Engineering and Architecture, 2011.
[21] L. Mao, *Smarandache Multi-Space Theory (I)—Algebraic Multi-Spaces*. New York, NY, USA: Cornell Univ., 2006.
[22] L. Mao, *Smarandache Multi-Space Theory (II)—Multi-Spaces on Graphs*. New York, NY, USA: Cornell Univ., 2006.
[23] L. Mao, *Smarandache Multi-Space Theory (III)—Map Geometries and Pseudo-Plane Geometries*. New York, NY, USA: Cornell Univ., 2006.
[24] L. Mao, *Proceedings of the First International Conference On Smarandache Multispace & Multistructures*. Beijing, China: The Educational Publisher Inc., 2013.
[25] F. Smarandache, "A unifying field in logics: Neutrosophic logic," *Multiple Valued Logic Int. J.*, vol. 8, no. 3, pp. 385–438, 2002.
[26] F. Hristea, *The Naïve Bayes Model for Unsupervised Word Sense Disambiguation. Aspects Concerning Feature Selection*. Amsterdam, The Netherlands: Springer, 2012.
[27] R. Sahin and M. Karabacak, "A multi attribute decision making method based on inclusion measure for interval neutrosophic sets," *Int. J. Eng. Appl. Sci.*, vol. 2, pp. 13–15, 1995.
[28] F. Xiao, "Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy," *Inf. Fusion*, vol. 46, pp. 23–32, Mar. 2019.
[29] P. Liu, Q. Khan, and T. Mahmood, "Some single-valued neutrosophic power muirhead mean operators and their application to group decision making," *J. Intell. Fuzzy Syst.*, vol. 37, no. 2, pp. 2515–2537, Sep. 2019.
[30] F. Smarandache, "Neutrosophic set—A generalization of the intuitionistic fuzzy set," in *Proc. 2006 IEEE Int. Conf. Granular Comput.*, Atlanta, GA, USA, Jun. 2006, pp. 38–42.
[31] F. Smarandache, *Plithogeny, Plithogenic Set, Logic, Probability, and Statistics*. Brussels, Belgium: Pons Publishing House, 2017.

MIHAELA COLHON received the Ph.D. degree in computer science from the Department of Computer Science, Faculty of Mathematics and Computer Science, The University of Piteşti, Romania, in 2009. She is currently an Associate Professor with the Department of Computer Science, University of Craiova, Romania. Her research fields are artificial intelligence, with specialization in knowledge representation, natural language processing (NLP), human language technologies (HLT), and computational statistics and data analysis with applications in NLP.

FLORENTIN SMARANDACHE received the M.Sc. degree in mathematics and computer science from the University of Craiova, Romania, and the Ph.D. degree in mathematics from the State University of Kishinev. He held a postdoctoral position in applied mathematics at the Okayama University of Sciences, Japan. He is currently a Professor of mathematics with the University of New Mexico, USA. He has been the Founder of neutrosophy (generalization of dialectics), neutrosophic set, logic, probability, and statistics, since 1995. He has published hundreds of articles on neutrosophic physics, superluminal and instantaneous physics, unmatter, absolute theory of relativity, redshift and blueshift due to the medium gradient and refraction index besides the Doppler effect, paradoxism, outerart, neutrosophy as a new branch of philosophy, Law of Included Multiple-Middle, multispace and multistructure, degree of dependence and independence between neutrosophic components, refined neutrosophic set, neutrosophic over-under-off-set, plithogenic set, neutrosophic triplet and duplet structures, quadruple neutrosophic structures, DSmT, and so on to many peer-reviewed international journals and many books. He presented articles and plenary lectures to many international conferences around the world. [http://fs.unm.edu/FlorentinSmarandache.htm]

DAN VALERIU VOINEA graduated from the Faculty of Economic Sciences and Business Administration, University of Craiova, Romania, the Faculty of History, Philosophy and Geography, University of Craiova, and the Ph.D. degree in sociology from the University of Craiova. He is currently a Senior Lecturer with the University of Craiova. He is also the author or coauthor of more than 30 scientific articles, including Web of science articles and of international seminars and conferences papers. He is a member of CCSCMOP's Scientific Committee.