Descubrimiento de nuevas palabras con polaridad afectiva a través de técnicas de inteligencia artificial general

Affective Polarity word discovering by means of Artificial General Intelligence techniques

Rafael del Hoyo, Isabelle Hupont, Francisco Lacueva
Instituto Tecnológico de Aragón
C/ María de Luna nº7
{rdelhoyo,ihupont,fjlacueva}@ita.es

Resumen: En este artículo se experimenta con la utilización de NARS como sistema de descubrimiento de palabras. y frases con contenido de opinión de forma dinámica.
Palabras clave: Sentimiento, NARS, AGI, Polaridad afectividad.

Abstract: In this article we make some experiments with NARS as words. and phrases with opinion dynamic content discovery..
Keywords: Sentiment, NARS, AGI, Polarity, Affective Computing

1 *General Introduction*

In general, there are two main approaches that aim at solving automatic opinion analysis of texts (a brief discussion on available approaches is given in Binali, 2009): statistical and semantic. Statistical approaches make use of data mining methods for opinion classification in texts, for instance, by using words and n-grams counts or by measuring distances for the automatic detection of the main words that provide affect. These data-mining techniques allow automatic opinion classification, but usually produce low classification results and need a wide number of texts for training. On the other hand, semantic approaches aim at classifying polarity in texts using information about a set of keywords, which implies the need of hand-crafted dictionaries (Osherenko, 2007, Shaikh, 2009) or looking for semantic ontologies like SenticNet (Cambria, 2012) and the incorporation of lexical analysis for detecting lexical modifiers of sense and negations in the phrase-level analysis.

Normally the second technique offer better results and accuracy. The generation of these dictionaries involves great handwork, being difficult to automate this process. There are possibilities of automatic generation of these dictionaries but the ambiguity of language, constant evolution of terms, phrases (Twitter is a good example) make necessary costly human supervision to maintain terms dictionaries with semantic opinion.

New techniques based in Artificial General Intelligence such as NARS (Wang, 2013), the Non-Axiomatic Reasoning System, aim to explain a large variety of cognitive phenomena with a unified theory. What makes NARS different from conventional reasoning systems is its ability to learn from its experience and to work with insufficient knowledge and resources, from a logic perspective. NARS attempts to uniformly explain and reproduce many cognitive facilities. We intend to use its logic and reasoning capabilities for modeling human language and thereby identifying the polarity of the new words. The main idea .of this experiment is to use a seed dictionary to look for new similar polarity words.

1.1 Affective polarity dictionary for seeding

One of the most followed psychological representations of affect considers emotions as a continuous 2D space whose dimensions are “evaluation” and “activation”. The “evaluation” dimension measures how a human feels, from positive to negative. The “activation” dimensions measures whether humans are more or less likely to take an action under an emotional state, from active to passive. Taking this affective 2D representation into account, the work of Cynthia Whissell provides a pair of <“activation”, “evaluation”> values (ranged from 1 to 3) to each one of the approx. 9000 affective words that compose her “Dictionary of Affect in Language” (DAL) (Whissell, 1989). The use of the DAL turns out interesting for the present work since it gives an “evaluation” value to each word, i.e. a polarity, depending on its affective contents. That way, the sentimental sensing in terms of affective contents can be extracted.

2 NARS introduction

NARS (Non-Axiomatic Reasoning System) (Wang, 2013) is a project aimed at the building of a general-purpose intelligent system, i.e., a "thinking machine" (also known as "AGI"), which follows the same principles as the human mind, and can solve problems in various domains. The design of NARS is based on the belief that the essence of intelligence is the principle of adapting to the environment while working with insufficient knowledge and resources. The logic part of NARS is NAL (Non-Axiomatic Logic), which is defined on a formal language called Narsese. To establish a solid semantic foundation for NAL, an Inheritance Logic, IL, is introduced first. One of its special property is that it is a term-oriented language, and belongs to the “term logic" school, rather than the “predicate logic". The basic component of Narsese is a term, an identifier that names a concept, which is a recognizable entity in the system's (internal and external) experience. An atomic term is just a sequence of characters from an alphabet. A Narsese statement relates a few terms to each other, and its basic form is an inheritance statement “S $\rightarrow$ P”, where ‘S’ is the subject term, ‘P’ the predicate term, and ‘$\rightarrow$’ the inheritance copula, which is a reflexive and transitive relation between two terms. The statement says that S is a specialization of P, and P is a generalization of S. All statements have 2 values that describe the uncertainty of the true. The first one is the frequency, proportion of positive evidence in all evidence, that is, f = w+ / w. The limit f, if exists, is the probability for the statement. In an open system, all frequency values may be changed by new evidence, and this is a major type of uncertainty, ignorance about the future frequency value. While frequency compares positive and negative evidence, a second measurement, confidence, can compare past and future evidence, in the same manner. Here the key idea is to only consider to a constant horizon in the future, that is, c = w/(w+k). A high confidence value means the statement is supported by more evidence, so less sensitive to new evidence. It does not mean that the statement is "closer to the reality", or the frequency is "closer to the true probability". A better description of the Narsese grammar can be obtained in Wang, 2013.

Cognitive linguistics differs from the traditional (computational or statistical) linguistics in its following hypotheses:

- language is not an autonomous cognitive faculty,
- grammar is conceptualization,
- knowledge of language emerge from language use.

The basic ideas of NLP in NARS (Wang, 2013b) can be summarized as the following:.

- lexical and grammatical knowledge are unified, and directly associated with semantic knowledge.
- linguistic knowledge can be derived from the system's experience in language usage (deduction, induction, abduction, and revision).

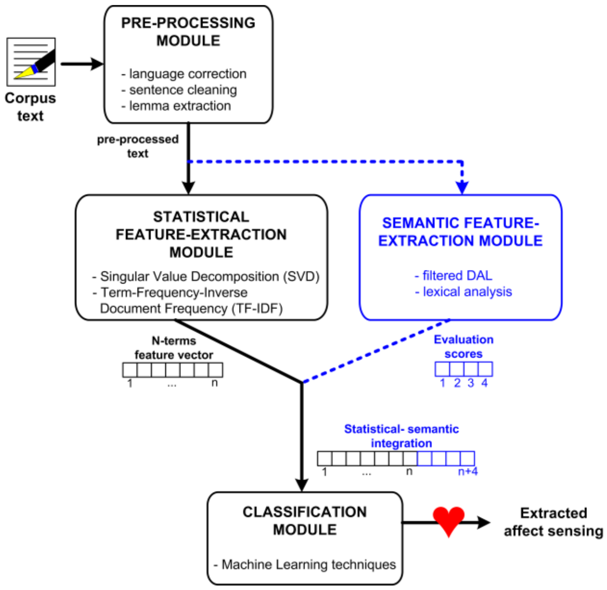

Fig. 1. **The hybrid system’s architecture overview used for experiments.**

3 Architecture Overview

The system for testing used for TASS 2013 is composed of four main blocks (see Figure 1), a better description can be read in (del-Hoyo, 2009). The pre-processing linguistic module is the first element block. The second element is in charge of statistics feature extraction, the third block provides the semantic information for the classification module. Finally, the last block is the classification module. In the following sub-sections, a general description of the modules is given.

4 Including Affective Information

To analyze the influence of "evaluation" dimension in texts' affect sensing, affective information is extracted from texts through the use of the Spanish translated DAL. Different syntactic categories of words can be found in the DAL: verbs, adjectives, pronouns, prepositions… However, many of them are context-dependent (e.g. the Spanish word "cara" means either "face" or "expensive", depending on the text context) and others are neutral in terms of affect. In order to minimize noise in the system and avoid possible confusions, a DAL filtering has been carried out for this work. Firstly, the words with neutral affective contents, i.e. with "evaluation" values ranging from 1.5 and 2.5, have been removed. Secondly, the rest of words have been divided into several word-lists, depending on their syntactic category and their "evaluation" value (considered " very negative" if lower than 1.25, "negative" if ranging from 1.25 and 1.5, "positive" if comprised between 2.5 and 2.75, and "very positive" if greater than 2.5). In particular, only the adjectives and verbs were retrieved from the DAL, since they are the groups of words with more contribution to emotional contents and more context-independent. Examples of built word-lists are: *<very positive adjectives>*, *<negative adjectives>*, *<positive verbs>*, etc. According to this process, the DAL is filtered and therefore the dictionary to work with is obtained.

The objective now is to provide each text with a general "evaluation" annotation. In order to achieve it, a set of structures or expert rules with different priorities are defined for the lexical analysis of texts. Those rules include the detection of lexical modifiers of meaning and negations in the phrase-level analysis. Examples of implemented rules are:

If there is a <positive adjective> then "evaluation" is <positive> (low priority rule)
If there is a <negative modifier> followed by a <very positive adjective> then "evaluation" is <very negative> (high priority rule)

Where *<negative modifier>* and *<positive adjective>* are lists of negation structures and positives adjectives, correspondingly.

Once the rules have been established, the texts can be processed. Four labels are defined: "very negative evaluation", "negative evaluation", "positive evaluation" and "very positive evaluation". Every time a structure of the aforementioned is found in the text, the score of the corresponding "evaluation" label is increased by one unit. That way, a global score for each "evaluation" label is obtained as an output for each processed text.

In order to include the extracted affective information into the statistical method, the obtained "evaluation" scores are added to the classification module. This is achieved by adding 4 new locations in the term vector build in the feature extraction module: one for each defined "evaluation" labels. In doing so, the semantic affective information is unified with the statistical method and a hybrid statistical-semantic system is achieved (see Figure 1).

5 Modelling Opinion with NARS

Several approaches can be generated to model opinion in Narsese. A simple initial approach has been selected, in order to obtain initial results. Every word in each Tweet has been cleaned (removed urls, nicks, brands..) and lemmatized. Only nouns of the tweets have been used, and the nouns have been labeled using DAL dictionary (very positive, positive, negative or very negative). Other nouns in the same twitters will be labeled with nouns closest to opinion words. In the following lines you can see example to model in Narsese opinion.

// amor term is a type of positive word.
<amor --> termino_positivo>. %1.00;0.90%

// analogy

<termino_muy_positivo <-> termino_positivo>. %0.90;0.60%
<termino_muy_negativo <-> termino_negativo>. %0.90;0.60%

// the words that are not in the DAL dictionary is in the same tweet with majority of positive words from DAL dictionary

<new_word --> cercano_termino_positivo>. %0.90;0.60%

// if any Word is close to positive terms then any Word is positive with probability 70% and confidence 60%
<<$x --> cercano_termino_positivo> ==> <$x --> termino_positivo>>. %0.70;0.60%

The inferred results are:

--- Level 39:
$0.39;0.03;0.32$ _@(T4-1) <new_word --> termino_positivo>. %0.90;0.35%

Highest logics can be generated for example:

// we define the relationship between new_word and affective words like amor.
<(,new_word,close_to_positive) --> amor>.*

// Si some_word is close to positive_word then some_word is positive
<<(,$x,close_to_positive)> ==> <$x --> termino_positivo>>. %0.90;0.60%*

<new_word --> termino_positivo>?
<?x --> termino_positivo>?

The Narsese allows us to use frequency and confidence in order to obtain the new words, the generation of the new words is confirmed with new tweets (confidence). Also other parameters like belief forgetting can improve the way to obtain new words.

6 Results and Conclusions

The TASS challenge was implemented in Java using two main tools Rapidminer and OpenNars. The OpenNars was used to discover new words in the affective dictionary. Rapidminer was used to implement the training process and text pre-processing part. The affective dictionaries was improved with a total of 21 new words (5 from very positive words, 3 positive words, 7 negative words, 6 very positive words). This new dictionary was created based on the evaluation twitter corpus, several of the words discovered was dictionary words bad formed (misspellings). The global system presented in section 3 was trained using different algorithms, like support vector machines, Multilayer Perceptron etc. The final classification algorithm and its configuration was selected using 10-Cross validation using twitter training data. The most important problem for discovering new words was current OpenNars performance issues for a big number of terms. The method used allows us to find new words dynamically in already existing dictionaries. Even though our approach is too much simple, it has allowed us to find small number of new words in the dictionary. Terms such as "jejej", or for example emoticons were found in this way. We are working to improve this approach by introducing more complex NAARS logic in order to model human language more accurately how is done in (Wang, 2013b) or using N-grams. Compared with the statistical approach of NLP, the NARS approach is different and new.

Bibliografía

Binali, H., Potdar, V., Wu, C.: A state of the art opinion mining and its application domains. IEEE International Conference on Industrial Technology, pp. 1--6 (2009).

Cambria E., and Hussain. Sentic Computing: Techniques, Tools, and Applications. Dordrecht, Netherlands: Springer, ISBN: 978-94-007-5069-2 (2012).

del-Hoyo, R., Hupont, I.. "Hybrid text affect sensing system for emotional language analysis." Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots. ACM, 2009.

Osherenko, A., André, E.: Lexical Affect Sensing: Are Affect Dictionaries Necessary to Analyze Affect? Proceedings of Affective Computing and Intelligent Interaction, Springer (2007)

Wang, P. Non-Axiomatic Logic: A Model of Intelligent Reasoning. World Scientific Publishing, 2013, ISBN: 9789814440271

Wang, P.: Natural language processing by reasoning and learning Proceedings of AGI-13, Pages 160-169, Beijing, China, August 2013

Shaikh, M., Prendinger, H., Ishizuka, M.: A Linguistic Interpretation of the OCC Emotion Model for Affect Sensing from Text. Affective Information Processing, Springer, pp. 45--73 (2009)

Whissell, C. M.: The Dictionary of Affect in Language. Emotion: Theory, Research and Experience: vol. 4, The Measurement of Emotions, R. Plutchik and H. Kellerman Eds. New York: Academic (1989)