BabelSenticNet: A Commonsense Reasoning Framework for Multilingual Sentiment Analysis

David Vilares
Language in the Information Society (LYS) Group
University of A Coruña
5 Rúa Almirante Lángara, 15011 A Coruña, Spain
david.vilares@udc.es

Haiyun Peng, Ranjan Satapathy, Erik Cambria
School of Computer Science and Engineering
Nanyang Technological University
50 Nanyang Ave, Singapore
{peng0065,satapathy.ranjan,cambria}@ntu.edu.sg

***Abstract*—SenticNet is a concept-level knowledge base used to develop commonsense reasoning algorithms for sentiment analysis tasks. One of the challenges that this resource must overcome is its lack of availability for languages aside from English. Prototype algorithms have been recently proposed to create non-English language concept-level knowledge databases, but they rely on a number of heterogeneous resources that complicate comparison, reproducibility and maintenance. This paper proposes an easy and replicable method to automatically generate SenticNet for a variety of languages, obtaining as a result BabelSenticNet. We use statistical machine translation tools to create a high coverage SenticNet version for the target language. We then introduce an algorithm to increase the robustness of the translated resources, relying on a mapping technique, based on WordNet and its multilingual versions. SenticNet versions for 40 languages have been made available. Human-based evaluation on languages belonging to different families, alphabets and cultures proves the robustness of the method and its potential for utility in future research on multilingual concept-level sentiment analysis.**

***Index Terms*—Multilingual Sentiment Analysis, Commonsense Knowledge**

I. Introduction

In recent years, aggregating and analyzing subjective information shared on social media has garnered the interest of many organizations and governments. Mining the public opinion about specific topics on social media platforms is crucial to many organizations for decision-making.

Sentiment analysis is the field of research focusing on the automatic extraction of opinions in human communication [1]. Many of such opinions, however, involve the presence of complex concepts that play a key role in the sentence and that current artificial intelligence (AI) techniques are unable to deal with. Such techniques usually rely on statistical methods based on co-occurrences of words or purely linguistic-based approaches, that are still far from being able to infer the cognitive and affective information associated with natural language, due to the limitations of the knowledge bases they rely on. For instance, `religious_experience` or `buy_christmas_present` are two expressions with subjective connotations that come from putting the words all together, and not from their individual terms. An AI that relies on mere objective connotations would perform poorly in discerning meaning and polarity of such expressions.

In the attempt to jump to the next curve of natural language processing (NLP) research [2], researchers must develop systems that have access to a significant amount of knowledge about the world and the domain of discourse. Only in doing so they can go beyond standard NLP algorithms, such as heavy domain-dependent applications or low coverage tools, which are becoming increasingly less efficient.

SenticNet [3] aims to tackle this challenge. It is a commonsense reasoning knowledge base intended for sentiment analysis, which contains semantic and affective information that connects various parts of extended common and commonsense knowledge representations to one another. The validity of this resource has been tested in a number of tasks [4], [5], [6]. However, developing non-English versions of SenticNet from scratch requires a number of resources that are usually unavailable in relation to the target language, such as ConceptNet [7] or AffectiveSpace [8].

This paper addresses the aforementioned problem. We propose BabelSenticNet, the first multilingual concept-level knowledge base for sentiment analysis: 40 languages have been made available for the research community both as an RDF/XML file[1] and as an API[2]. An initial version of SenticNet for the target language is created through statistical machine translation (SMT). We then increase the robustness of the translated knowledge base by using an over-mapping technique. This is done through relying on WordNet and the available version of WordNet for the target language, so as to match SenticNet concepts with WordNet synsets and find their translations in the target WordNet.

The remainder of the paper is organized as follows: Section II reviews related research on sentiment analysis; Section III describes our approach for BabelSenticNet; Section IV shows a human-based evaluation on a subset of BabelSenticNet, proving that the cross-linguistic approach is able to keep a large number of semantically related concepts, all while taking into account languages belonging to different families, alphabets and cultures; finally, in Section V we draw conclusions and outline future research.

[1] http://sentic.net/babelsenticnet.zip
[2] http://sentic.net/api

978-1-5386-9276-9/18/$31.00

II. Related research

Sentiment analysis techniques can be broadly categorized into symbolic and sub-symbolic AI approaches: the former include the use of lexicons [9], ontologies [10], and semantic networks [11] to encode the polarity associated with words and multiword expressions; the latter consist of supervised [12], semi-supervised [13] and unsupervised [14] machine learning techniques that perform sentiment classification based on word co-occurrence frequencies. Among these, the most popular approaches are supervised: given a set of labeled data, a variety of features is extracted for each sample, and then used to feed a supervised classifier, relying on features such as n-grams [15], part-of-speech tags [16] or enriched generalized dependency triplets [17].

Pang and Lee [18] showed that developing methods which only analyze the subjective portions of an opinion helps improve the performance of machine-learning models for sentiment analysis. Marchetti-Bowick and Chambers [19] introduced the concept of distant supervision, which consists in training a sentiment analysis model using subjective data collected from external sources, such as a social media network or a forum. Mohammad et al. [20] collected data from Twitter, for the purpose of creating subjective lexica from it, using a point-wise mutual information based method. The research team then used the lexica as a starting point to extract features for the supervised classifier. More recent approaches are based on deep neural networks and generative adversarial networks [21], [22]. Severyn and Moschitti [23] used a convolutional neural network, using pre-trained word embeddings and distant supervision, obtaining one of the best performing systems at SemEval 2015 [24].

Another approach is the lexicon-based method, which combines subjective lexica with rule-based systems to compute the final polarity of an opinion. Thelwall et al. [25] presented SentiStrength, an unsupervised approach for dual-score sentiment analysis on English short-texts. They considered the characteristic phenomena of this style of texts; such as the poor grammatical quality of the micro-texts, the replication of characters or the overuse of capital letters. Later works [26], [27] described a similar method for long reviews, which rely on semantic orientation values and morphological-based rules to handle relevant linguistic phenomena for the purpose in question; such as negation, intensification, *irrealis* or subordinate adversative clauses. Vilares et al. [28] proposed a syntactic version of their approach, attaining significant improvements on the same datasets, and demonstrating the potential utilization of dependency parsing for polarity classification tasks.

A. Concept-level sentiment analysis

Existing AI algorithms for sentiment analysis are far from being able to infer the cognitive and affective information associated with natural language. To this end, the SenticNet initiative has been developing resources for enabling sentiment analysis at the semantic, rather than syntactic, level since 2010: SenticNet 1 simply associated polarity scores with almost 6,000 ConceptNet concepts; in addition to polarity, SenticNet 2 also assigned semantics and sentics to commonsense concepts and extended the breadth of the knowledge base to about 13,000 entries; SenticNet 3 broadened the spectrum of the semantic network to 30,000 concepts; SenticNet 4 introduced the concept of semantic primitives to further extended the knowledge base to 50,000 entries; finally, SenticNet 5 reached 100,000 commonsense concepts by employing recurrent neural networks to infer primitives by lexical substitution.

The correct way to use SenticNet is in concomitance with *sentic patterns* [29], a collection of syntax-based rules that describe how the concepts and relevant linguistic phenomena should interact in a sentence. They show how sentiment should flow from concept to concept, based on the dependency relation of the input sentence and, hence, generate a binary (positive or negative) polarity value reflecting the feeling of the speaker (Fig. 1).

B. Multilingual sentiment analysis resources

There are a number of recent works on the definition of language-specific methods for opinion mining in a wide variety of languages, including Arabic [30], Chinese [31], French [32], German [33], Hindi [34], Italian [35], Japanese [36], Russian [34], Spanish [28] and Thai [37]. One of the problems researchers face when dealing with languages aside from English, is the lack of sentiment dictionaries [38]. A current research direction is the automatic or semi-automatic generation of large, non-English resources which are then applied to reasonably performing methods for sentiment analysis [39].

SentiStrength lexical resources [25] were automatically translated to make the system available for a variety of languages. For some of them, such as Spanish, the resources were later improved by leveraging on additional lexica, attaining a significant improvement over the original foreign language, e.g., Spanish - SentiStrength [40]. Hogenboom et al. [41] proposed projecting sentiment scores from English SentiWordNet to the Dutch version, by exploiting the relations between English WordNet [42] and its Dutch counterpart [43].

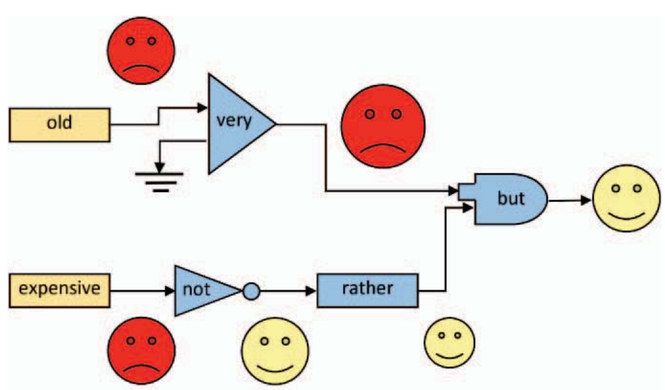

Fig. 1. Visualization of how sentic patterns model the sentiment data flow for the opinion target ‘car’ in the sentence “the car is old but rather not expensive”.

Ghorbel and Jacot [32] translated English SentiWordNet entries intro French, and found that even if the translation is correct, it does happen that two parallel words do not always share the same semantic orientation across both languages, simply due to a difference in common usage. To solve this issue, Volkova et al. [44] proposed the use of crowdsourcing: the researchers applied bootstrapping so as to learn sentiment lexicons for English, Spanish and Russian language texts from Twitter streams. The process starts with a set of seed words obtained from existing corpus in the case of the English language writings. For the other two languages, English seed terms were translated using bilingual dictionaries. Chen and Skiena [45] proposed a method for building high-quality sentiment lexicons for 136 languages by integrating a variety of linguistic resources to produce a knowledge graph. By appropriately propagating from seed words, they constructed sentiment lexicons for each language of the graph. Their experiments showed a 95.7% agreement with published lexicons and a coverage of 45.2%. However, to the best of our knowledge, there are no available resources for multilingual sentiment analysis at the concept level, reinforcing the novelty of our work and its impact for the NLP community.

III. BabelSenticNet

Let $C = \{c_1, c_2, ...c_n\}$ be the set of concepts, where c_i is a single or multi-word expression representing an affective concept (e.g., `beautiful` or `buy_christmas_present`), let $S_i = \{s_{i1}, s_{i2}, ..., s_{i5}\}$ be the set of *semantics*, i.e., the set of concepts $s_{ij} \in C$, semantically related to the concept c_i, with $i \in |C|$ and $j \in [1, 5]$, and let A={*pleasantness*, *attention*, *sensitivity*, *aptitude*, *polarity*} be a set of affective attributes where $\forall a \in A \rightarrow \quad a \in \mathbb{R}$ and $a \in [-1, 1]$; let $H \subset$ {*joyful*, *admirable*, *disgusting*, *sad*, *interesting*, *angry*} with $|H| = 2$ be a set of emotions arisen by a concept, SenticNet is defined as an unidirectional, cyclic, semantic graph G={c, S, A, H}. An English language version of SenticNet (from now G_{en}) is currently available to the research community in RDF/XML format[3]. Figure 2 illustrates the semantic subgraph for the concept `buy_christmas_present`.

This level of representation of knowledge has potential advantages for sentiment analysis tasks. In the aforementioned example, it would be easy to associate the concept `buy_christmas_present` with concepts such as `annual_celebration`, `fight_inflation` or `winter_time`, providing a useful context for sentiment analysis algorithms to make decisions not merely based on the sentence's content.

Given G_{en} and a set of target languages, T, the aim is to obtain a set of translated graphs, G_t with $t \in T$. Some first attempts to define methodologies for translating concept-level knowledge bases, such as SenticNet, have been recently proposed. In 2014, Xia et al. [46] presented an approach to build a Chinese version of SenticNet.

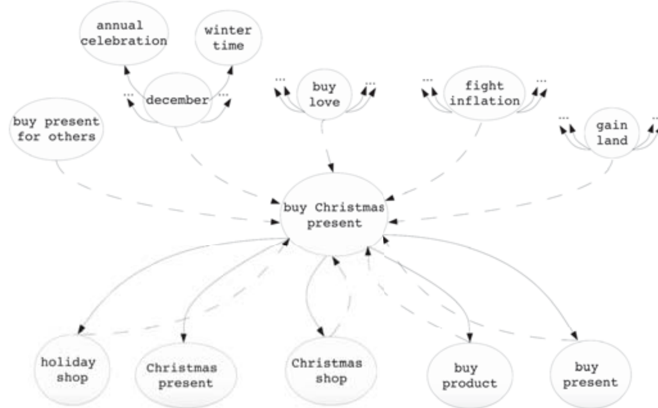

Fig. 2. Semantic subgraph for the concept `buy_christmas_present`. Continuous lines represent the semantics of a concept. Dashed lines illustrates some concepts that have `buy_christmas_present` as a semantic.

They used web dictionaries as the starting point to find the counterpart of an English concept, c_i and then to obtain a set of semantics S_i. The approach, although potentially applicable to other languages, requires a large quantity of resources in order to obtain a non-English version of SenticNet, which are often unavailable.

In this paper, the translation process of G_{en} only involves the set of concepts C, and not the affective or emotion attributes of the graph (i.e., given a concept $c_i \in C$ -¿ $A_{(i,en)}$=$A_{(i,t)}$) and $H_{(i,en)}$=$H_{(i,t)}$. In the same vein, Xia et al. [46] suggested that some concepts ought to have changed affective attributes when they are translated from one language to another, due to cultural differences. For example, the concept `dragon` is usually associated with `luck` in the Chinese language, but the same connotation is not necessary true for Indoeuropean languages.

Carrying out this process for a large number of target languages is, however, both costly and time consuming, and it is beyond the scope of this paper. However, existing cross-linguistic models for sentiment analysis utilizing SMT techniques have shown their ability to achieve state-of-the-art results [27]. Additionally, the human-based evaluation reported in Section IV reinforces the practical utility of our cross-linguistic approach as a strong baseline.

A. Creating the non-English versions of SenticNet

We first created an automatically translated version for the various considered languages from the English version of SenticNet. To do so, we used the Bing translator API[4], a state-of-the-art SMT tool often used on NLP tasks [47], [48], [49]. We observed that these translated versions, although possessing a high coverage, presented some weaknesses:

- Some concept translations were wrong. Both single and multiword expressions were translated from English SenticNet, but no additional context was provided, which is a drawback for ambiguous words.

[3]http://sentic.net/senticnet-5.0.zip

[4]http://microsoft.com/en-us/translator/translatorapi.aspx

- A number of English language concepts were translated to a unique concept in the target language, decreasing the coverage for the corresponding non-English version of SenticNet.
- Some concepts simply could not be translated.

To overcome these issues, we introduced an algorithm that improves the robustness of the translated versions of SenticNet.

B. Increasing the robustness with WordNet

WordNet [42] is a lexical database for English, where each word is associated with one or more *synsets*. A synset consists of an identifier that represents a unique meaning. A word can have one or more synsets (e.g., `love` can be in reference to the verb or the noun, which will be represented as two different synsets) and one synset can be related to one or more words (e.g., `beautiful` and `good-looking` can be synonyms).

There are also a number of non-English versions for WordNet available, such as the Spanish, Dutch, Italian [43] or Chinese WordNet [50]. The mapping is not always complete, i.e., not all the English terms have a corresponding representation in all the target languages. However, a common feature among the different language versions of WordNet is that synsets are shared. Thus, given an expression and its synset, it is possible to find an ideal translation in the target language. We exploit this advantage in order to increase the robustness of our automatically translated versions of SenticNet.

However, some issues must be taken into account. In some cases, corresponding WordNet terms for some languages were not available (e.g., Hindi). In other situations, there are WordNet versions for languages for which machine translation was unavailable (e.g., Galician or Vasque). For the first group, the automatically translated version is released without additional modifications. For the second group, we used the corresponding WordNet version to create a reduced version of SenticNet, by solely using the mapping technique. Although coverage is lower in the latter case, this allows us to provide the first concept-level knowledge base for a number of languages that totally lack resources of this type.

There are two main challenges in improving the robustness of BabelSenticNet: (a) given that a word can have homonyms, choosing the right synset is important for the purposes of increasing the accuracy of the automatically translated version and (b) a word present in SenticNet does not necessary have a direct corresponding term in English WordNet. To solve the first challenge, several strategies were considered:

a) *Select a random synset* (if exists).

b) *Select the first entry returned by WordNet* (if exists).

We also took into account two additional heuristics that consider the set of semantics, S_i, given the concept c_i, that can even be used in case (b):

1) *Select the most frequent synset in semantics*: Let $synset_x \in Synsets_{c_i}$ be a synset for the concept c_i with $x \in |Synsets_{c_i}|$, and let $Synsets_{S_i}$ be a list composed of all the synsets that can be extracted from each of the semantics in S_i. This method will select $\text{argmax}_{synset_x} f(synset_x, Synsets_{S_i})$ where the function f counts the number of occurrences of the $synset_x$ in $Synsets_{S_i}$.
2) *Select the most semantically related synset*: Computed as $\text{argmax}_{synset_{ix}} g(synset_{ci_x}, Synsets_{S_i})$ where g is a function to measure the taxonomy similarity between two synsets. There are a number of metrics to measuring synset similarities in terms of the taxonomy hypernym/hyponym (e.g., Wu-Palmeror or Resnik Similarity), but for the purposes of this work, we rely on the Leacock-Chodorow (LCH) similarity, after it was observed in preliminary experiments that different metrics produced similar results for the purpose at hand.

Strategy	Single word	Multiword
Random	0.65	0.50
First	0.80	0.80
Most frequent in semantics	0.90	0.85
LCH	0.90	0.85

TABLE I
ACCURACY OF THE PROPOSED SYNSET SELECTOR STRATEGIES IN MATCHING SENTICNET CONFLICTIVE CONCEPTS TO A WORDNET SYNSET

To verify the validity of the metrics, we took 20 single word and 20 multiword concepts that were manually annotated as a valid or invalid synset for the concept, given their semantics and the synsets' definition.

It is important to note that more than one synset can be valid. Table I shows the accuracy for the different proposed metrics. Based on our evaluation, we chose the synset selector strategy (2), although strategy (1) also achieved good results. The algorithm to predict the most likely synset given a concept and its semantics is shown in Algorithm 1, where $+$ is the operation that connects two lists.

Algorithm 1 Predict the most likely synset given a concept and its semantics

procedure MOST_LIKELY_SYNSET(c_i, S_i)
    synsets $\leftarrow$ `WordNet`(c_i)
    if $|synsets| = 0$ **then return** $\{\}$
    if $|synsets| = 1$ **then return** synsets[0]
    if $|synsets| > 1$ **then**
        $semanticsynsets \leftarrow []$
        for s_{ij} in S_i **do**
            $semanticsynsets +$ `WordNet` (s_{ij})
    $most\_similar\_synset \leftarrow \varnothing$
    $similarity\_best\_synset \leftarrow -\infty$
    for $synset$ in $synsets$ **do**
        for ss in $semanticsynsets$ **do**
            if $g(synset, ss) > similarity\_best\_synset$ **then**
                $most\_similar\_synset \leftarrow ss$
                $similarity\_best\_synset \leftarrow g(synset, ss)$
    return $most\_similar\_synset$

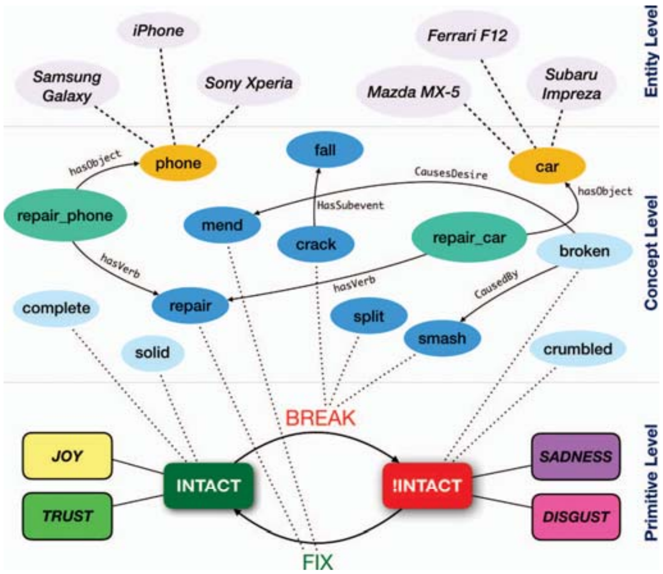

Fig. 3. Hierarchical structure of SenticNet. Translation of multiword expressions can be highly enhanced if the conceptual primitive associated to each word is known.

In the future, we plan to enrich this algorithm with additional information coming from SenticNet's hierarchical structure, i.e., we will not only look at connections at the Concept Level but also at the Primitive Level (Fig. 3). This will help great deal in choosing the right translation for multiword expressions, e.g., by substituting a term for which translation is missing with a germane term associated to the same conceptual primitive.

IV. Experiments

We relied on a human-based evaluation to determine the quality of our cross-linguistic approach. Due to the number of SenticNet versions released together with this paper, an exhaustive evaluation was unfeasible. Hence, we only selected five languages, namely: Spanish, Portuguese, Italian, Hindi and Chinese. The reason behind this choice is that these languages include some of the most spoken languages on the Web (besides English) and that they are from different linguistic families, alphabets and cultures, which thus gives us a clearer idea of the accuracy of this approach under different conditions. For each language, we focused on the first 500 concepts and their semantics.

We asked two fluent annotators of the target language to answer 'yes' or 'no' for each semantic, in response to the question "*Is the semantic actually semantically related to its concept?*". For example, given the concept `dog`, it is clear that concepts such as `pet` or `mammal` are semantically related to it. In a more subtle way, concepts such as `go_for_a_walk` or `play_in_the_park` are also related, since it is not uncommon to do these tasks with a pet. On the flip side, it seems clear that other concepts such as `computer` or `car_factory` do not possess any clear relation in respect to it.

On the one hand, it is important to note that it is unlikely that an annotator answers 'yes' wrongly because the ability to confirm that a semantic and a concept are actually semantically related is due to acquired knowledge or experience in the field. On the other hand, other annotators might not find a relation between a semantic and its concept due to the lack of experience or contextual knowledge, since SenticNet also contains some highly specialized or technical concepts (referring to fields such as biology, medicine, and chemistry).

A. Metrics

Let R_a be the list of length N containing the response tuples (s_i, r), where r is the response of the annotator a to the semantic s_i, with $i \in [1, N]$; we are using three different metrics to measure the quality of the translated SenticNet:

$$At\ least\ one\ yes = \frac{|R_{a_1}(yes) \cup R_{a_2}(yes) \cup R_{a_3}(yes)|}{N} \quad (1)$$

where $Ra_j(yes)$ represents the semantics for which annotator j marked 'yes'.

$$Majority\ yes = \frac{\sum_i^N majority(s_i, yes)}{N} \quad (2)$$

where $majority(s_i, yes)$ returns 1 if two or more annotators labeled the semantic s_i with 'yes' and 0 otherwise.

$$Total\ agreement = \frac{|R_{a_1}(yes) \cap R_{a_2}(yes) \cap R_{a_3}(yes)|}{N} \quad (3)$$

B. Results

Table II illustrates the results of the human-based evaluation. The percentage of semantics marked as semantically related to its concept, for at least one annotator, varies from 70.7% (Chinese) to 95.8% (Hindi). The results indicate that a high number of the translated semantics are semantically related to the corresponding translated concept, considering different family languages, which reinforces the robustness and practical utilities of the approach for future multilingual concept-level sentiment analysis.

It is important to note that a concept marked as 'no' by the annotators does not necessary implicate that the cross-linguistic approach did not work for that concept or semantic, nor that the translated SenticNet has a lower quality, as we briefly comment in Section IV-C.

	Spanish	**Chinese**	**Hindi**
Annotator 1	0.580	0.680	0.893
Annotator 2	0.420	0.390	0.865
Annotator 3	0.690	0.707	0.684
At least one 'yes'	0.754	0.707	0.958
Majority of 'yes'	0.594	0.649	0.8848
Total agreement	0.338	0.358	0.598

TABLE II
RESULTS (%) OF THE HUMAN-BASED EVALUATION FOR THE SPANISH, CHINESE AND HINDI VERSIONS OF SENTICNET.

Concept	Semantic 1	Semantic 2	Semantic 3	Semantic 4	Semantic 5
abandon	leane alone	smelly foot	unattractive	lose team	leave out
renunciar	dejar	mal olor de pies[2]	poco atractivo[1]	peder el equipo[1]	omitir[1]
abduction	reaper	intolerance	danger	illiteracy	kidnap
secuestro	~~usuario~~[2]	intolerancia a la[2]	peligro	analfabestismo[2]	secuestrar
able read	good sense of humor	good eyesight	taken seriously	complement	much need
capaz de leer	buen sentido del humor[2]	buena vista	tomado en serio[2]	complemento[2]	necesita mucho[2]

TABLE III
EXAMPLES OF CONCEPTS AND THEIR SEMANTICS. SOME OF THEM WERE DETERMINED, BY AT LEAST ONE OF THE ANNOTATORS, TO BE NOT SEMANTICALLY RELATED. IF THE TRANSLATED TEXT IS WRONGLY TRANSLATED, THE TEXT IS SHOWN IN STRIKETHROUGH FONT. THE SUPERSCRIPTS INDICATE THE NUMBER OF ANNOTATORS WHO FOUND THE SEMANTIC TO BE SEMANTICALLY UNRELATED TO ITS CONCEPT.

The percentage of the semantic relations identified by two or more annotators varies from 58% (Spanish) to 89.3% (Hindi). The results suggest that for the three studied languages, in the majority of cases the users found the semantic to be related to the corresponding concept. Finally, we also report for how many concepts all the annotators agreed on the translated semantic to be related to the concept, varying from 33.8% (Spanish) to 59.8% (Hindi).

We can conclude that, depending on the language, between 30% and 60% of the semantics related to the concept would be considered correct by most people, between 60% and 90% of the semantics are considered to be semantically related by a significant number of people and between 70% and 95% of the semantics can be only recognized by very few people.

As SenticNet is in the end a sentiment lexicon, it is also important to prove that its translated versions keep a high percentage of semantics where the polarity of a concept based on their semantics is coherent. To prove this, we took Spanish (the language for which obtained the worst results in terms of semantics related to the concept) and we asked a native speaker to determine for every of the 500 concepts if the polarity that was directly assigned from English was right, based on the concept and the semantic translations.

We obtained that around 64.4% of the assigned polarities were correct. It is important to note that some of the concepts marked as incoherent were terms where the polarity was highly dependent on the context (e.g., medical terms such as `paracetamol` or `adrenaline` were considered as negative and positive, respectively).

C. Error analysis

Although the number of semantic relations identified by at least one annotator is high, there are still a significant number of semantics that were found to be unrelated to their concepts. We discovered that the main reasons why annotators did not relate a semantic to its concept were:

- *A concept or semantic could not be translated to the target language from English SenticNet and the English form was kept*. There was however a small number of cases that were marked as semantically related to the concept (e.g., anglicisms widely used in the target language).
- *A concept or semantic was wrongly translated*, even after running the over mapping algorithm.
- *The semantic was correctly translated from English SenticNet, but some of the annotators still did not find a reliable semantic relation to the corresponding concept*. We show some examples in Table III comparing the English and Spanish versions of SenticNet, where many of the translations were right, but were marked as not being semantically related by some, or all, of the annotators.

V. CONCLUSION

This paper presented BabelSenticNet, the first concept-level knowledge base for multilingual sentiment analysis. The resource is available for free in 40 languages. To build it, we combined statistical machine translation with an over mapping approach based on English WordNet and its multilingual versions.

The proposed method is low in cost and not time consuming. The human-based evaluation reinforces the robustness of the method across different languages, alphabets and cultures and opens a new path for future research in multilingual concept-level sentiment analysis. In the future, we plan to enrich the proposed method with additional information coming from SenticNet's hierarchical structure, e.g., by exploiting conceptual primitives to improve the translation of multiword expressions.

REFERENCES

[1] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.

[2] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," *IEEE Computational Intelligence Magazine*, vol. 9, no. 2, pp. 48–57, 2014.

[3] E. Cambria, S. Poria, D. Hazarika, and K. Kwok, "SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings," in *AAAI*, 2018, pp. 1795–1802.

[4] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowledge-Based Systems*, vol. 108, pp. 42–49, 2016.

[5] M. Araújo, P. Gonçalves, M. Cha, and F. Benevenuto, "Ifeel: A system that compares and combines sentiment analysis methods," in *Proceedings of the companion publication of the 23rd international conference on World wide web companion*. International World Wide Web Conferences Steering Committee, 2014, pp. 75–78.

[6] E. Cambria, A. Hussain, T. Durrani, C. Havasi, C. Eckl, and J. Munro, "Sentic computing for patient centered applications," in *IEEE ICSP*, 2010, pp. 1279–1282.

[7] R. Speer and C. Havasi, "Conceptnet 5: A large semantic network for relational knowledge," in *Theory and Applications of Natural Language Processing*, E. Hovy, M. Johnson, and G. Hirst, Eds. Springer, 2012, ch. 6.

[8] E. Cambria, J. Fu, F. Bisio, and S. Poria, "AffectiveSpace 2: Enabling affective intuition for concept-level sentiment analysis," in *AAAI*, 2015, pp. 508–514.

[9] A. Bandhakavi, N. Wiratunga, S. Massie, and P. Deepak, "Lexicon generation for emotion analysis of text," *IEEE Intelligent Systems*, vol. 32, no. 1, pp. 102–108, 2017.

[10] M. Dragoni, S. Poria, and E. Cambria, "OntoSenticNet: A commonsense ontology for sentiment analysis," *IEEE Intelligent Systems*, vol. 33, no. 3, pp. 77–85, 2018.

[11] S. Poria, A. Gelbukh, E. Cambria, P. Yang, A. Hussain, and T. Durrani, "Merging senticnet and wordnet-affect emotion lists for sentiment analysis," in *Signal Processing (ICSP), 2012 IEEE 11th International Conference on*, vol. 2. IEEE, 2012, pp. 1251–1255.

[12] L. Oneto, F. Bisio, E. Cambria, and D. Anguita, "Statistical learning theory and ELM for big social data analysis," *IEEE Computational Intelligence Magazine*, vol. 11, no. 3, pp. 45–55, 2016.

[13] A. Hussain and E. Cambria, "Semi-supervised learning for big social data analysis," *Neurocomputing*, vol. 275, pp. 1662–1673, 2018.

[14] Y. Li, Q. Pan, T. Yang, S. Wang, J. Tang, and E. Cambria, "Learning word representations for sentiment analysis," *Cognitive Computation*, vol. 9, no. 6, pp. 843–851, 2017.

[15] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques," in *Proceedings of EMNLP*, 2002, pp. 79–86.

[16] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA), May 2010.

[17] D. Vilares, M. A. Alonso, and C. Gómez-Rodríguez, "On the usefulness of lexical and syntactic processing in polarity classification of twitter messages," *Journal of the Association for Information Science Science and Technology*, vol. 66, no. 9, pp. 1799–1816, 2015.

[18] L. Pang, Bo and Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2004, pp. 271–278.

[19] M. Marchetti-Bowick and N. Chambers, "Learning for microblogs with distant supervision: Political forecasting with twitter," in *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. Avignon, France: ACL, Apr. 2012, pp. 603–612.

[20] S. M. Mohammad, S. Kiritchenko, and X. Zhu, "Nrc-canada: Building the state-of-the-art in sentiment analysis of tweets," in *SemEval*, 2013.

[21] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[22] Y. Li, Q. Pan, S. Wang, T. Yang, and E. Cambria, "A generative model for category text generation," *Information Sciences*, vol. 450, pp. 301–315, 2018.

[23] A. Severyn and A. Moschitti, "Unitn: Training deep convolutional neural network for twitter sentiment classification," in *SemEval*, 2015, pp. 464–469.

[24] S. Rosenthal, P. Nakov, S. Kiritchenko, S. M. Mohammad, A. Ritter, and V. Stoyanov, "Semeval-2015 task 10: Sentiment analysis in twitter," in *SemEval*, 2015.

[25] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the social web," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 1, pp. 163–173, 2012.

[26] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.

[27] J. Brooke, M. Tofiloski, and M. Taboada, "Cross-linguistic sentiment analysis: From english to spanish," in *Proceedings of the International Conference RANLP-2009*. Borovets, Bulgaria: ACL, 2009, pp. 50–54.

[28] D. Vilares, M. A. Alonso, and C. Gómez-Rodríguez, "A syntactic approach for opinion mining on spanish reviews," *Natural Language Engineering*, vol. 21, no. 01, pp. 139–163, 2015.

[29] S. Poria, E. Cambria, A. Gelbukh, F. Bisio, and A. Hussain, "Sentiment data flow analysis by means of dynamic linguistic patterns," *IEEE Computational Intelligence Magazine*, vol. 10, no. 4, pp. 26–36, 2015.

[30] M. A. Ibrahim and N. Salim, "Opinion analysis for twitter and arabic tweets: A systematic literature review," *Journal of Theoretical and Applied Information Technology*, vol. 56, no. 2, pp. 338–348, Oct. 2013.

[31] H. Peng, Y. Ma, Y. Li, and E. Cambria, "Learning multi-grained aspect target sequence for chinese sentiment analysis," *Knowledge-Based Systems*, vol. 148, pp. 167–176, 2018.

[32] H. Ghorbel and D. Jacot, "Sentiment analysis of french movie reviews," in *Advances in Distributed Agent-Based Retrieval Tools*, ser. Studies in Computational Intelligence, V. Pallotta, A. Soro, and E. Vargiu, Eds. Berlin and Heidelberg: Springer, 2011, vol. 361, pp. 97–108.

[33] T. Scholz and S. Conrad, "Linguistic sentiment features for newspaper opinion mining," in *NLDB*, ser. Lecture Notes in Computer Science. Springer, 2013, vol. 7934, pp. 272–277.

[34] N. Medagoda, S. Shanmuganathan, and J. Whalley, "A comparative analysis of opinion mining and sentiment classification in non-english languages," in *Proceedings of ICTER 2013, International Conference on Advances in ICT for Emerging Regions*. Colombo, Sri Lanka: IEEE, Dec. 2013.

[35] F. Neri, C. Aliprandi, F. Capeci, M. Cuadros, and T. By, "Sentiment analysis on social media," in *Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE, Aug. 2012, pp. 951–958.

[36] Y. Arakawa, A. Kameda, A. Aizawa, and T. Suzuki, "Adding twitter-specific features to stylistic features for classifying tweets by user type and number of retweets," *Journal of the Association for Information Science and Technology*, vol. 65, no. 7, pp. 1416–1423, 2014.

[37] P. Inrak and S. Sinthupinyo, "Applying latent semantic analysis to classify emotions in thai text," in *2nd International Conference on Computer Engineering and Technology (ICCET)*, vol. 6. Chengdu: IEEE, Apr. 2010, pp. 450–454.

[38] S. L. Lo, E. Cambria, R. Chiong, and D. Cornforth, "Multilingual sentiment analysis: From formal to informal and scarce resource languages," *Artificial Intelligence Review*, vol. 48, no. 4, pp. 499–527, 2017.

[39] R. Steinberger, "A survey of methods to easy the development of highly multilingual text mining applications," *Language Resources and Evaluation*, vol. 46, pp. 155–176, 2012.

[40] D. Vilares, M. Thelwall, and M. A. Alonso, "The megaphone of the people? spanish sentistrength for real-time analysis of political tweets," *Journal of Information Science*, vol. 41, no. 6, pp. 799–813, 2015.

[41] A. Hogenboom, B. Heerschop, F. Frasincar, U. Kaymak, and F. de Jong, "Multi-lingual support for lexicon-based sentiment analysis guided by semantics," *Decision Support Systems*, vol. 62, pp. 43–53, 2014.

[42] G. A. Miller, "Wordnet: a lexical database for english," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[43] P. Vossen, Ed., *EuroWordNet. A Multilingual Database with Lexical Semantic Networks*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1998.

[44] S. Volkova, T. Wilson, and D. Yarowsky, "Exploring sentiment in social media: Bootstrapping subjectivity clues from multilingual twitter streams," in *ACL 2013. 51st Annual Meeting of the Association for Computational Linguistics Proceedings of the Conference. Volume 2: Short Papers*. Sofia, Bulgaria: ACL, Aug. 2013, pp. 505–510.

[45] Y. Chen and S. Skiena, "Building sentiment lexicons for all major languages," in *The 52nd Annual Meeting of the Association for Computational Linguistics. Proceedings of the Conference. Volume 2: Short Papers. ACL 2014*. Baltimore: ACL, Jun. 2014, pp. 383–389.

[46] Y. Xia, X. Li, E. Cambria, and A. Hussain, "A localization toolkit for SenticNet," in *ICDMW*, 2014, pp. 403–408.

[47] J. J. Castillo, "Using machine translation systems to expand a corpus in textual entailment," in *Advances in Natural Language Processing*. Springer, 2010, pp. 97–102.

[48] P. Pakray, S. Neogi, S. Bandyopadhyay, and A. Gelbukh, "A textual entailment system using web based machine translation system," in *NTCIR*, 2011.

[49] B. S. Dhakar, S. K. Sinha, and K. K. Pandey, "A survey of translation quality of english to hindi online translation systems (google and bing)," *International Journal of Scientific and Research Publications*, p. 313, 2013.

[50] J. Huang, S.-K. Hsieh, J.-F. Hong, Y.-Z. Chen, I.-L. Su, Y.-X. Chen, and S.-W. Huang, "Chinese wordnet: Design, implementation, and application of an infrastructure for cross-lingual knowledge processing," *Journal of Chinese Information Processing*, vol. 24, no. 2, pp. 14–23, 2010.