Multi-lingual Support for Lexicon-Based Sentiment Analysis Guided by Semantics

Alexander Hogenboom, Bas Heerschop, Flavius Frasincar*

Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, the Netherlands

Uzay Kaymak

Eindhoven University of Technology, P.O. Box 513, NL-5600 MB Eindhoven, the Netherlands

Franciska de Jong

Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, the Netherlands

Universiteit Twente, P.O. Box 217, NL-7500 AE Enschede, the Netherlands

Abstract

Many sentiment analysis methods rely on sentiment lexicons, containing words and their associated sentiment, and are tailored to one specific language. Yet, the ever-growing amount of data in different languages on the Web renders multi-lingual support increasingly important. In this paper, we assess various methods for supporting an additional target language in lexicon-based sentiment analysis. As a baseline, we automatically translate text into a reference language for which a sentiment lexicon is available, and subsequently analyze the translated text. Second, we consider mapping sentiment scores from a semantically enabled sentiment lexicon in the reference language to a new target sentiment lexicon, by traversing relations between language-specific semantic lexicons. Last, we consider creating a target sentiment lexicon by propagating sentiment of seed words in a semantic lexicon for the target language. When extending sentiment analysis from English to Dutch, mapping sentiment across languages by exploiting relations between semantic lexicons yields a significant performance improvement over the baseline of about 29% in terms of accuracy and macro-level F_1 on our data. Propagating sentiment in language-specific semantic lexicons can outperform the baseline with up to about 47%, depending on the seed set of sentiment-carrying words. This indicates that sentiment is not only linked to word meanings, but tends to have a language-specific dimension as well.

Keywords: Multi-lingual sentiment analysis, semantics, lexicon, machine translation, map, propagation

1. Introduction

In today's complex, globalizing markets, information monitoring tools are of paramount importance for decision makers. Such tools help decision makers in identifying issues and patterns that matter, as well as in tracking and predicting emerging events. Traditional decision support systems typically provide support for decisions by accurately deriving actionable knowledge from structured data, whereas the extraction of useful information from unstructured data like natural language text still poses important challenges [1]. Recent advances in tools for information extraction have been primarily focused on retrieving explicit pieces of information from natural language text on different levels of granularity [2]. State-of-the-art information monitoring and extraction tools enable us to identify entities like companies, products, or brands in text, and to subsequently extract more complex concepts, such as

*Corresponding author; tel: +31 (0)10 408 1340; fax: +31 (0)10 408 9162

Email addresses: hogenboom@ese.eur.nl (Alexander Hogenboom), basheerschop@gmail.com (Bas Heerschop), frasincar@ese.eur.nl (Flavius Frasincar), u.kaymak@ieee.org (Uzay Kaymak), f.m.g.dejong@utwente.nl (Franciska de Jong)

events in which these entities play various roles [3]. Recent research endeavors additionally explore how to perform such information extraction tasks on a multitude of heterogeneous sources in an ever-changing environment [4, 5, 6].

However, latent pieces of information can be extracted from natural language text as well. For instance, recent work has made it possible to detect the distinct topics that people discuss in their (on-line) conversations [7, 8]. Yet, for many application scenarios, it is not so much the entities, events, or topics that people discuss per se, but rather people's sentiment with respect to these subjects that provides decision makers with valuable information. This is reflected by the recent surge in research interest in sentiment analysis for decision support [1, 9, 10, 11].

Sentiment analysis techniques can support decision making in a multitude of scenarios. For instance, sentiment analysis can help organizations pinpoint the effect of specific issues on customer perceptions, thus helping these organizations respond with appropriate marketing and public relations strategies [12]. Furthermore, consumer sentiment has been demonstrated to have a significant impact on stock ratings [13, 14] and sales [15, 16]. Thus, accurate sentiment analysis methods are crucial for supporting decision making in these fields. Additionally, tracking of stakeholders' sentiment is important for decision making in economic systems [17], financial markets [18], politics [19], organizations [20], and reputation management [21].

Real-world decision support systems typically consist of four logical components, i.e., a Knowledge Management System (KMS), a Model Management System (MMS), a Database Management System (DMS), and a User Interface System (UIS) [22]. Each of these logical components is used to monitor and regulate the flow of crucial information in order to support decision making in an organization. Data managed by the DMS can be transformed into actionable knowledge in the KMS, with the MMS controlling how the obtained knowledge is used in models in order to support decision making, and the UIS taking care of the interaction with the end user of the system. In order to utilize sentiment-based information in decision support systems, the DMS should be enriched with (user-generated) sentiment-carrying content that has been crawled from the Web. Furthermore, the KMS should be able to represent indicators of identified sentiment with respect to a topic of interest. Additionally, the MMS should allow for the incorporation of sentiment-based information in the decision making process. Last, the UIS should provide dashboards with relevant information that enables decision makers to act upon arising issues in a timely manner.

One of the key open issues that must be resolved in order to be able to exploit the full potential of sentiment analysis in real-life decision support systems is that these systems must be able to deal with textual data in various languages [1]. Such data is available in vast amounts, as recent developments on the Web enable users to produce an ever-growing amount of virtual utterances of opinions or sentiment through, e.g., messages on Twitter, blogs, or reviews, in any language of their preference.

The analysis of sentiment in the overwhelming amount of available multi-lingual textual data is challenging at best. This challenge can be addressed by means of automated sentiment analysis techniques, focusing on determining the polarity of natural language text. Typical approaches involve scanning a text for cues signalling its polarity, e.g., (parts of) words or other (latent) features of natural language text. Lexicon-based sentiment analysis methods have gained (renewed) attention in recent work [23, 24, 25, 26, 27, 28, 29], not in the least because their performance has been shown to be robust across domains and texts [30]. Such methods essentially rely on lexical resources containing words and their associated sentiment, i.e., sentiment lexicons, and their nature allows for intuitive ways of accounting for structural or semantic aspects of text in sentiment analysis [26, 31].

Many existing lexicon-based sentiment analysis approaches are tailored to one specific language – typically English. However, in order for automated sentiment analysis to be useful for decision makers in today's complex, globalizing markets, automated sentiment analysis tools need to be able to support multiple languages rather than English only. Therefore, we explore how we can analyze sentiment in another language – i.e., Dutch – for which we have nothing more but some lexical and syntactical parsing tools, a semantic lexical resource, and a handful of positive and negative sample words.

A good starting point is SentiWordNet [32, 33], as recent research has proven this large (semantic) sentiment lexicon for English, generated by means of machine learning techniques, to be rather effective when used for analyzing sentiment in texts published in our reference language, i.e., English [34].

As a first step, one could consider translating texts from a target language, i.e., Dutch, to our reference language, i.e., English, in order to be able to subsequently utilize the well-established SentiWordNet sentiment lexicon for the reference language in the sentiment analysis process.

However, as subjectivity is associated with word meanings rather than words [35], literal translation of texts to a reference language in order to benefit from the available sentiment lexicon for the reference language may be suboptimal in automated sentiment analysis of texts in another language. As an alternative, we therefore propose to map the sentiment from the reference sentiment lexicon to a sentiment lexicon for the target language, by means of traversing relations between large language-specific semantic lexical resources, thus accounting for word meanings rather than lexical representations. Additionally, we consider an approach that involves propagating sentiment from a seed set of words in a language-specific semantic lexical resource for each considered language separately, in order to generate language-specific sentiment lexicons which can subsequently be used in language-specific sentiment analysis methods.

The main contribution of our work lies in our novel sentiment mapping method, which exploits relations between language-specific semantic lexicons in order to construct a sentiment lexicon for a target language. We compare the effectiveness of this method with that of an existing machine-translation approach and a method that focuses on semantic relations within, rather than across languages. We thus aim to provide insight in the importance of semantics for multi-lingual sentiment analysis.

The remainder of the paper is organized as follows. In Section 2, we discuss related work on (multi-lingual) sentiment analysis and the semantic lexicons that may be exploited in this process. We then elaborate on our framework for assessing our considered methods for dealing with another language in sentiment analysis in Section 3. Our findings are discussed in Section 4. We conclude and provide directions for future work in Section 5.

2. Related Work

Today's abundance of user-generated content has resulted in a surge of research interest in systems that are able to deal with opinions and sentiment, as explicit information on user opinions is often hard to find, confusing, or overwhelming [36]. Many language-specific sentiment analysis approaches exist, whereas the exploration of how to support multiple languages when analyzing sentiment has only just begun.

2.1. Sentiment Analysis

The roots of sentiment analysis are in fields like natural language processing, computational linguistics, and text mining. The main objective of most sentiment analysis approaches is to extract subjective information from natural language text. Most work focuses on determining the overall polarity of words, sentences, text segments, or documents [36]. This task is commonly approached as a binary classification problem, in which a text is to be classified as either positive or negative. However, this task may be approached as a ternary classification problem as well, by introducing a third class of neutral documents. An alternative to such sentiment classification approaches is the determination of a degree of positivity or negativity of natural language text in order to produce, e.g., rankings of positive and negative documents [37, 38].

Many state-of-the-art approaches to sentiment classification tasks rely on machine learning techniques [36, 39]. On the other hand, some approaches exploit (generic) sentiment lexicons when determining the subjectivity or polarity of natural language text. Both approaches may be combined in hybrid methods as well [29].

In machine learning sentiment analysis methods, natural language text is typically modeled by means of a bag-of-words vector representation, denoting an unordered collection of words occurring in this text. In order to be able to, e.g., distinguish pieces of text from one another in terms of their associated polarity class, machine learning methods typically aim to find and exploit patterns in the vector representations of these texts. In such vector representations, a binary encoding scheme, indicating the presence of specific words, has proven to be effective [39] as well as to outperform frequency-based encoding [40]. Vectors may also contain features other than words, e.g., parts of words, word groups, or features representing semantic distinctions between words [41]. Features represented in vectors may be weighted as well [42].

Lexicon-based methods account for the semantic orientation of individual words in a text by matching these words with a list of words and their associated sentiment scores, i.e., a sentiment lexicon.

The text's overall semantic orientation is then determined by aggregating (e.g., summing) the individual word scores, as retrieved from the sentiment lexicon. Hybrid approaches may realize the aggregation through a machine learning process as well [29]. In this sentiment scoring process, other aspects of content may be taken into account as well, such as negation [27, 43], intensification [28], or the rhetorical roles of text segments [26, 31].

As they often incorporate deep linguistic analysis into the sentiment detection procedures [26], lexicon-based sentiment analysis methods tend to sacrifice computational efficiency for a classification accuracy which is typically inferior to the classification accuracy of machine learning methods in specific domains for which machine learning approaches can be trained and optimized [30]. However, lexicon-based approaches have an attractive advantage over machine learning methods in that they have a more robust performance across domains and texts [30]. Additionally, lexicon-based approaches enable deep linguistic analysis to be incorporated into the sentiment analysis process [26] which, if fine-tuned, can improve the classification accuracy. Moreover, lexicon-based sentiment scoring approaches are essentially rule-based approaches, which can inherently provide insight into the motivation for the classification of the conveyed sentiment. Last, lexicon-based approaches can be generalized relatively easily to other languages by using dictionaries [35].

2.2. Multi-lingual Sentiment Analysis

Today's sentiment analysis systems must deal with an abundance of multi-lingual sentiment-carrying user-generated content. As different approaches are required for distinct languages [44], existing work does not typically focus on devising a single sentiment analysis approach for multiple languages, but rather on analyzing the sentiment conveyed by documents in selected languages, mainly by means of applying sentiment analysis techniques tailored to each specific language. Existing work is primarily focused on how to devise sentiment analysis methods for other languages with minimal effort, without sacrificing too much accuracy. Rather than constructing new frameworks for languages other than the reference language [44, 45, 46, 47, 48], recent work focuses on using machine translation techniques in order to be able to re-use many existing tools when performing automated sentiment analysis on multi-lingual natural language content.

Sentiment analysis of machine-translated texts may seem a rather ineffective approach, as machine translation typically fails to correctly translate substantial amounts of text and moreover tends to reduce well-formed texts to sentence fragments. Nevertheless, recent work on sentiment analysis of news messages in nine languages demonstrates that the accuracy of sentiment classification on machine-translated text is largely independent of the quality of the machine translator used (i.e., the translator does not necessarily have to produce well-formed texts) and that sentiment analysis of texts that have been translated into English is rather consistent across languages, after normalizing sentiment scores in order to allow for meaningful cross-cultural comparisons [49].

Other work suggests that in some cases, sentiment analysis of machine-translated texts can yield even better results than sentiment analysis of the original texts [50]. This appears to be the case especially when the original language is not easily interpreted by state-of-the-art natural language processing tools. For instance, in [50], the authors use a Chinese framework for classifying the sentiment of Chinese reviews, and an English framework for classifying the sentiment of Chinese reviews that have been translated into English. The results indicate that sentiment analysis of the translated texts outperforms sentiment analysis of the original texts. An ensemble of both methods further improves the performance.

Machine translation can be utilized in another way as well in order to facilitate automated sentiment analysis in multiple languages. Rather than performing sentiment analysis on machine-translated texts, many researchers focus on automatically generating sentiment lexicons by means of machine translation. A common approach is to automatically translate an existing sentiment lexicon [35], and, possibly, to subsequently propagate the sentiment scores to semantically related words [51]. An alternative approach, which has been shown to outperform machine translation of sentiment lexicons, is to automatically generate a sentiment lexicon from a collection of (automatically) translated and annotated texts [35, 52, 53, 54]. However, research suggests that the subjectivity of most of the words in sentiment lexicons is lost in translation – subjectivity appears to be a property associated not with words, but with word meanings [35]. Semantic lexicons can be used in order to address this issue.

2.3. Semantic Lexicons

A widely used on-line (semantic) lexical resource is WordNet [55], the design of which has been inspired by psycholinguistic theories of human lexical memory. WordNet is organized into sets of cognitive synonyms – synsets – which can be differentiated based on their Part-of-Speech (POS) type. Each WordNet synset expresses a distinct concept and is linked to other synsets through different kinds of relations, such as synonymy, antonymy, hyponymy, or meronymy. The need for such a lexical reference system has arisen as conventional dictionaries do not usually capture such semantic relations. Conventional dictionaries use lexicographical sorting for words for human users' convenience. Conversely, WordNet has been designed to be used under program control and enables the distinction between different word forms and word meanings.

SentiWordNet [32, 33] is a lexical resource in which each WordNet synset is associated with three numerical scores, quantifying its associated sentiment. These scores describe how objective, positive, and negative the terms contained in a synset are. An ensemble of eight ternary classifiers has been used to classify each synset as either objective, positive, or negative, based on a vector representation of the associated description of the synset. The overall objectivity, positivity, and negativity scores for a synset have then been determined by the (normalized) proportion of classifiers that assigned the corresponding labels to the synset.

The availability of semantic lexical resources is not limited to the English language. For instance, EuroWordNet [56] has been developed as a collection of semantic lexicons for several European languages, including English, Dutch, Italian, and Spanish. For each supported language, a semantic lexicon has been created, with a structure similar to the structure of WordNet. Additionally, EuroWordNet has been designed in such a way that the language-specific semantic lexicons are linked to one another through WordNet, such that each English synset is associated with its equivalents in the languages included in EuroWordNet.

For Dutch, i.e., the language considered in our current work as an alternative to our English reference language, a more extensive semantic lexicon has been developed on top of EuroWordNet as well. In DutchWordNet (Cornetto) [57], the Dutch part of EuroWordNet has been enriched with information from the Referentie Bestand Nederland (RBN), which is a lexical database for Dutch, containing information on orthography, morphology, syntax, semantics, pragmatics, and combinatorics.

Language-specific semantic lexical resources and their interlinkage through semantic lexical resources such as WordNet can facilitate new approaches for extending an existing lexicon-based sentiment analysis approach from one language to another. The semantic relations between language-specific semantic lexicons could be exploited in order to propagate a sentiment lexicon from one language to another, while preserving semantics. Alternatively, sentiment scores for a seed set of words could be propagated through a language-specific semantic lexicon in order to generate language-specific sentiment lexicons [34, 58, 59, 60]. As both types of approaches account for semantics, they may compensate for the drawbacks of existing machine translation methods for multi-lingual sentiment analysis.

3. Framework

In order to investigate how lexicon-based sentiment analysis can be extended from our reference language, i.e., English, to another language, i.e., Dutch, we first need a lexicon-based sentiment analysis framework for the reference language. This framework can then serve as a starting point for an extension to another language.

3.1. Polarity Classification

Building upon our previous work [34], we use a binary polarity classifier that classifies documents as either positive or negative based on the aggregated sentiment scores for individual words, as retrieved from a semantic sentiment lexicon such as SentiWordNet. For an arbitrary synset, we compute a single sentiment score based on its objectivity, positivity, and negativity scores (all positive real numbers which sum to 1), by subtracting the negativity score from the positivity score, thus obtaining a real number in the interval $[-1, 1]$, representing sentiment scores in the range from negative to positive, respectively.

In our polarity classification process, detailed in Algorithm 1, documents are first split into sentences and words. Then, each word's POS type, lemma, and word sense are determined in order to subsequently retrieve its sentiment score from the sentiment lexicon. For the word sense disambiguation process, we use a Lesk-based algorithm for WordNet [61], as described in our previous work [34].

For each word in a sentence, the algorithm essentially selects the word sense that is semantically most similar to the words in the context, i.e., the other words in a sentence.

After retrieving all word-level sentiment scores from the sentiment lexicon, the sentiment score ζ_d of a document d is computed by summing the sentiment scores ζ_t of each non-stopword t in each sentence s of the document, i.e.,

$$\zeta_d = \sum_{s \in d} \sum_{t \in s} \zeta_t. \tag{1}$$

The resulting document-level sentiment score is subsequently used in order to classify the document's polarity class c_d as either positive (1) or negative (-1), i.e.,

$$c_d = \begin{cases} 1 & \text{if } (\zeta_d - \epsilon) \geq 0, \\ -1 & \text{if } (\zeta_d - \epsilon) < 0, \end{cases} \tag{2}$$

with ϵ representing an offset that corrects for a possible bias towards positivity in sentiment scores. Such a bias may be caused by people's tendency to write negative texts with rather positive words [30]. Following existing work [30], we calculate the offset ϵ on a training set as

$$\epsilon = 0.5 \left(\frac{\sum_{d \in P} \zeta_d}{|\Phi|} + \frac{\sum_{d \in N} \zeta_d}{|N|} \right), \tag{3}$$

with Φ denoting the subset of positive documents in the training set, and N denoting the subset of negative documents in the training set.

Algorithm 1: Classifying a document's polarity.

input : A document d and an offset ϵ
output: The polarity classification c_d of document d

```
ζd = 0;
foreach s ∈ d do
    foreach t ∈ s do
        pos = findPOS(t, s);
        lemma = findLemma(t, pos);
        sense = findWordSense(t, s, pos);
        ζt = getWordScore(lemma, sense, pos);
        ζd = ζd + ζt;
    end
end
cd = 1;
if (ζd − ε) < 0 then
    cd = −1;
end
return cd;
```

Our sentiment analysis framework has been developed for classifying the polarity of English documents. As such, in order to be able to classify the polarity of documents written in another language, the latter documents could be automatically translated into the reference language, such that they can be analyzed by means of the sentiment analysis framework for the reference language. Thus, our existing English sentiment analysis framework can be used for classifying the polarity of Dutch documents without needing to develop any new natural language processing components other than a machine-translation component.

However, the concepts of our framework can be used for polarity classification in Dutch as well, if lexical and syntactical parsing tools for identifying sentences, words, POS, and lemmas are available for Dutch, as well as a semantic lexical resource for the Dutch language. The latter semantic lexical resource can be used for word sense disambiguation, as well as for constructing a Dutch sentiment lexicon that can be used in a sentiment analysis framework with components tailored to the Dutch language.

Our framework (visualized in Fig. 1) supports two of such alternatives to the machine translation approach. First, we consider traversing the relations between language-specific semantic lexicons in order to map the existing sentiment lexicon for the English reference language to a new sentiment lexicon for the Dutch target language. This method is detailed in Section 3.2. Second, we consider propagating sentiment within language-specific semantic lexical resources, as described in Section 3.3.

3.2. Traversing Relations between Language-Specific Semantic Lexical Resources

The valuable information contained in the sentiment lexicon of an existing sentiment analysis approach for the reference language can be utilized in another language when it is used to generate a sentiment lexicon for the target language. This may be done by (automatically) translating an existing sentiment lexicon from the reference language into the target language [35, 51]. However, as subjectivity tends to be associated with word meanings rather than words [35], we propose a novel method of translating a sentiment lexicon from a reference language to a target language, while taking into account the semantics of the words in the sentiment lexicons. To this end, we exploit language-specific semantic lexical resources and their interrelations.

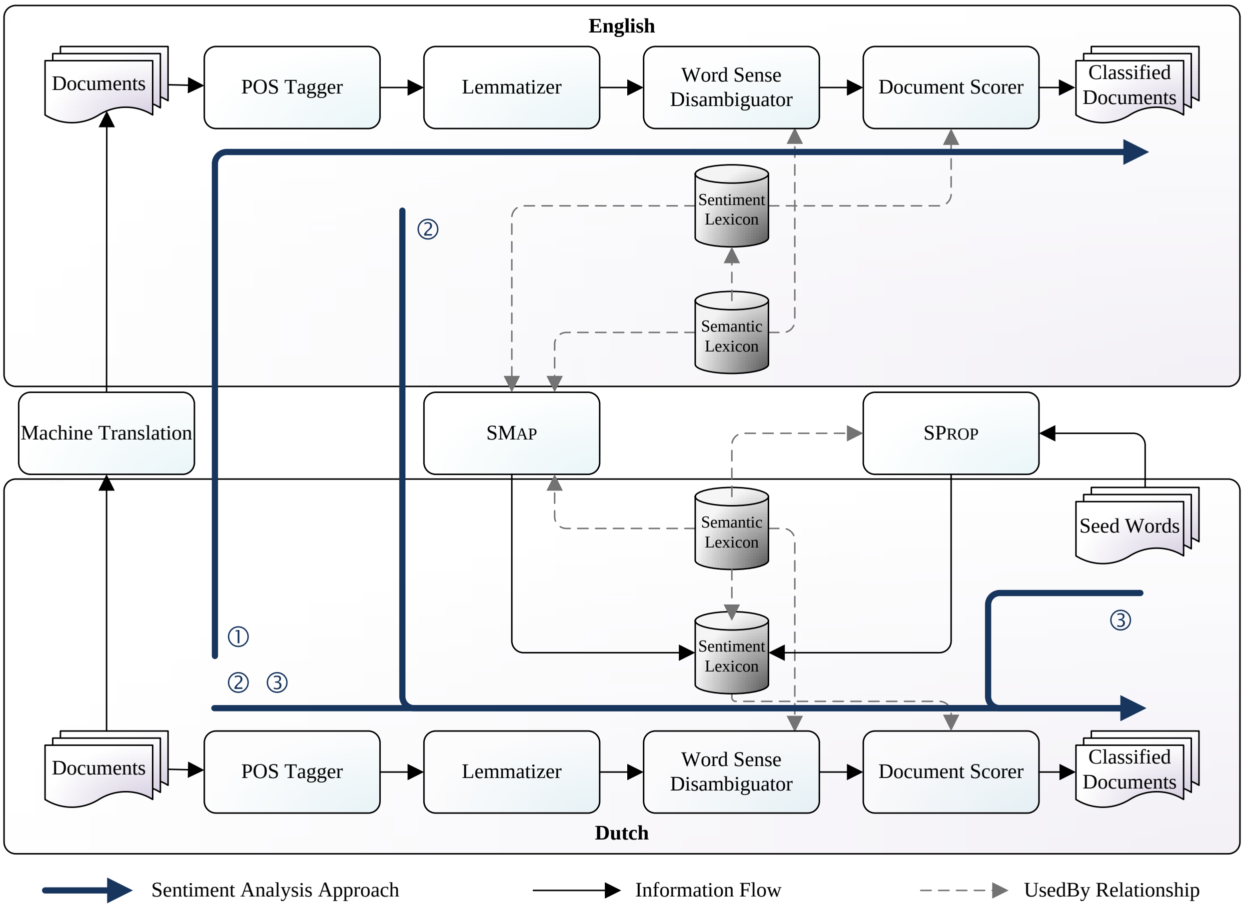

Figure 1: Our employed sentiment analysis framework, with language-specific components for both English and Dutch. Our three considered approaches of using these components in order to analyze the sentiment of Dutch documents are marked with bold arrows. Approach ① is to translate our Dutch documents into English and to subsequently use the available existing English sentiment analysis components. The alternative approaches ② and ③ involve analyzing the sentiment of our Dutch documents by means of Dutch language-specific components while exploiting a sentiment lexicon that has been constructed based on either an existing English sentiment lexicon (②), or seed sets of Dutch sentiment-carrying words (③).

In our novel cross-lingual sentiment score mapping method SMAP (see Fig. 2), we assume an existing sentiment lexicon for the reference language to be linked to a semantic lexical resource with meaningfully related words and concepts (synsets). Provided that a mapping exists between this semantic lexicon and an equivalent semantic lexicon for another language, the sentiment from the reference sentiment lexicon can be mapped to a new sentiment lexicon for the target language by traversing the associated relations between the semantic lexicons of both respective languages.

For example, for our reference language (English) and target language (Dutch), the English SentiWordNet sentiment lexicon can be used as a starting point for our proposed cross-lingual sentiment score mapping procedure. SentiWordNet contains sentiment scores for all synsets in the WordNet semantic lexicon. Additionally, a mapping exists between WordNet and its Dutch equivalent DutchWordNet (Cornetto). By exploiting these relations, SentiWordNet sentiment scores associated with English WordNet synsets can be projected onto equivalent Dutch synsets in DutchWordNet (Cornetto), thus yielding a Dutch sentiment lexicon.

In order to propagate sentiment scores associated with synsets through language-specific semantic lexical resources, we first map the reference sentiment lexicon's synsets to the reference semantic lexicon. Subsequently, we map the synsets in the semantic lexicon for the new language to their equivalent synsets in the reference semantic lexicon.

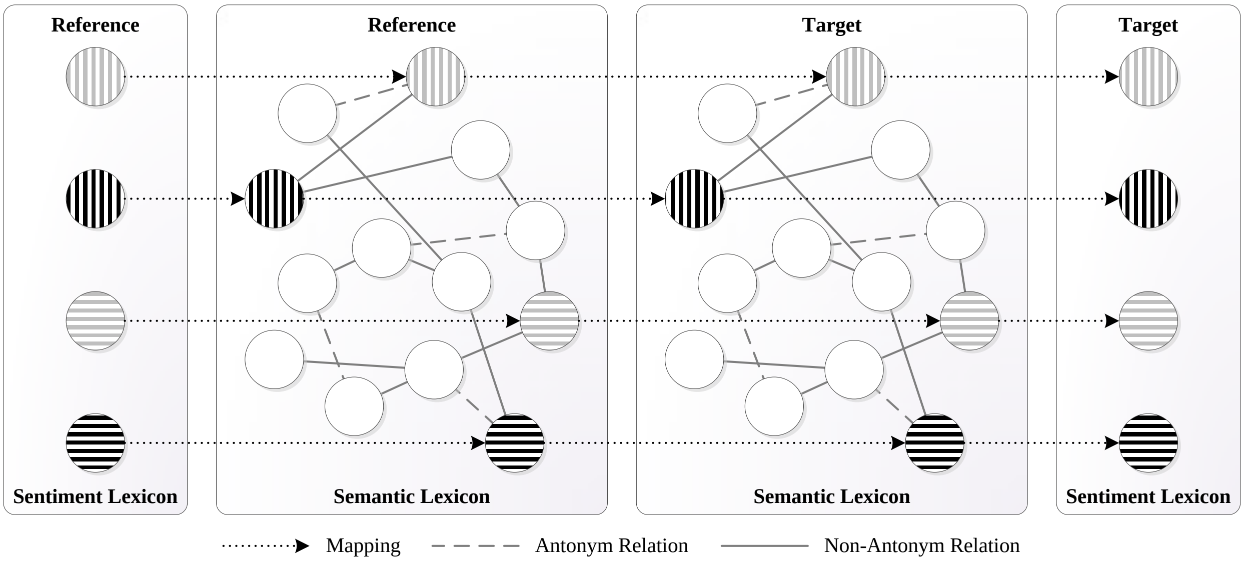

Figure 2: Our novel method for mapping sentiment scores from a reference language to a target language. Positive words and synsets are marked with vertical stripes, whereas negative words and synsets are marked with horizontal stripes. Others are left blank. Darker shading implies stronger sentiment.

Then, for each synset in the reference sentiment lexicon, we use these mappings to assign the associated reference sentiment score to the equivalent synsets and their synonyms in the semantic lexicon for the target language. The result is saved in the new sentiment lexicon for the target language. This process is further detailed in Algorithm 2.

Algorithm 2: Sentiment propagation through relations between semantic lexical resources.

input : The reference sentiment lexicon S^*, the reference semantic lexicon L^*, and the semantic lexicon for the target language L'
output: The sentiment lexicon for the target language S'

$S' = \emptyset$;
foreach $sentiWord^* \in S^*$ **do**
  $synset^* =$ `getSynset(`$sentiWord$, S^*`)`;
  $synset' =$ `mapSynsetFromTo(`$synset^*$, L^*, L'`)`;
  if $synset' \neq \emptyset$ **then**
    $pos =$ `getPos(`$synset^*$, L^*`)`;
    $\zeta =$ `getScore(`$synset^*$, S^*`)`;
    $synonyms =$ `getSynonyms(`$synset'$, L'`)`;
    foreach $t \in synonyms$ **do**
      $lemma =$ `getLemma(`t, L'`)`;
      $sense =$ `getWordSense(`t, L'`)`;
      $S' = \{S', \{synset', lemma, sense, pos, \zeta\}\}$;
    end
  end
end
return S';

3.3. Sentiment Propagation within Language-Specific Semantic Lexical Resources

When creating a new sentiment lexicon for a new target language, one could also consider not to use the reference sentiment lexicon as a starting point, as the sentiment associated with words or word meanings may have a cultural dimension. Instead, one could consider creating a new sentiment lexicon for the target language by means of an approach involving propagating the sentiment of a small seed set of words to words which are semantically related [34, 58, 59, 60].

In our sentiment propagation method SProp (detailed in Algorithms 3 and 4 and visualized in Fig. 3), semantic relations in a language-specific semantic lexicon are traversed for each seed word. Examples of such semantic relations are hyponymy (type-of relations), synonymy, and antonymy. In the sentiment propagation process, each encountered word t is stored with a sentiment score ζ_t, based on the score ξ of the seed word, a diminishing factor δ, and the number of steps k (with a maximum of K) between the seed word and t, i.e.,

$$\zeta_t = \xi\tau\delta^k, \quad \tau \in \{-1, 1\}, \quad k \in \{1, \ldots, K\}, \\ -1 \leq \xi \leq 1, \quad 0 < \delta < 1, \qquad (4)$$

with τ indicating whether to invert (-1) the score, i.e., when traversing antonym relations, or not (1).

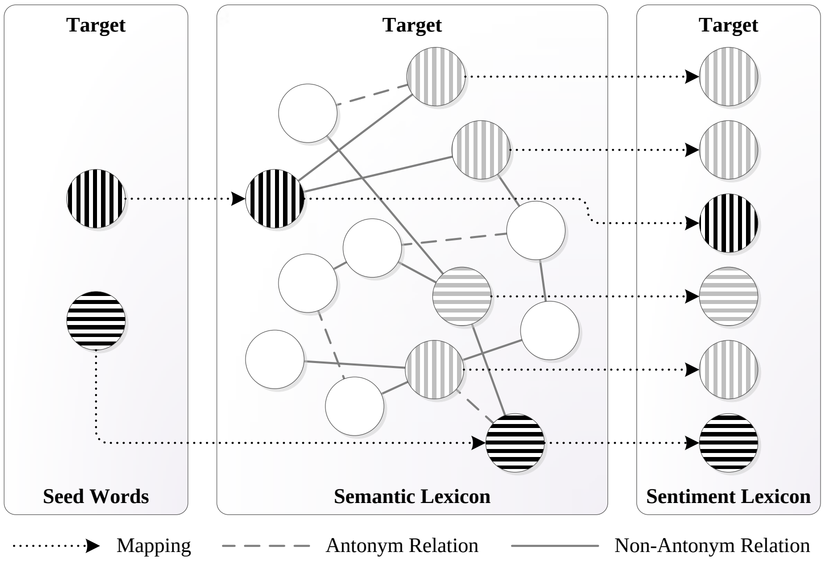

Figure 3: Our proposed method for propagating the sentiment of a set of seed words through the semantic lexicon of a target language. The sentiment lexicon thus generated in one propagation step is visualized. Positive words and synsets are marked with vertical stripes, whereas negative words and synsets are marked with horizontal stripes. Others are left blank. Darker shading implies stronger sentiment.

In each iteration of our algorithm, the computed sentiment score ζ_t for the current word t is propagated to the words in its directly related synsets. While doing so, with each next traversed semantic relation, the propagated sentiment is further diminished. As a result, words that are semantically more closely related to a seed word obtain a higher absolute sentiment score than those with a more indirect semantic relation to a seed word. If a word is encountered multiple times when propagating the sentiment associated with seed words, this word is assigned the score obtained from the shortest path between the word and any of the seeds, because we assume that the shorter the path, the more accurate the sentiment can be determined.

4. Evaluation

An evaluation of the performance of the methods proposed in Section 3 can provide insight in how lexicon-based sentiment analysis can best be extended from our reference language, i.e., English, to another language, i.e., Dutch. This evaluation can help understand the importance of semantic relations – both across and within languages – for multi-lingual sentiment analysis. In this section, we present our experimental set-up and discuss our experimental results.

4.1. Experimental Setup

We focus on 600 positive and 600 negative opinionated Dutch documents on 40 distinct topics, crawled from Dutch review Web sites, forums, and blogs. The documents have been classified by three human annotators, until they reached full consensus. On this corpus, we assess the performance of our considered methods by means of the 10-fold cross-validated overall sentiment classification accuracy and the macro-level F_1-score. We assess the statistical significance of performance differences by means of a paired two-sample one-tailed t-test.

The implementation of our sentiment classification framework has been done in C#.Net. We have built upon our existing framework for classifying the sentiment of English documents [34], which classifies sentiment as described in Section 3.1. We have constructed a similar implementation for sentiment classification of Dutch documents, which is an extension of the English implementation by means of the translation and sentiment propagation methods discussed in Section 3.

For classifying the sentiment of English text, our implementation uses regular expressions in order to split the text into words. POS tagging is done with a SharpNLP [62] POS tagger. Lemmatization and word sense disambiguation is performed by means of the C# WordNet.Net [63] WordNet API. The

Algorithm 3: Propagating sentiment in a language-specific semantic lexical resource.

input : The target semantic lexicon L', a list *seeds* of the words to propagate, their associated scores ξ, the maximum number of iterations K, and a diminishing factor δ

output: A sentiment lexicon S' containing all propagated words with their computed sentiment scores

```
syn = getSynsets(L');
S' = ∅;
foreach t ∈ seeds do
    ξ = score(t);
    S' = propWord(syn, S', t, ξ, δ, 1, K); // Alg. 4
end
return S';
```

Algorithm 4: Propagating a single word's sentiment in a lexical resource (propWord).

input : All *synsets* in the semantic lexicon for the target language, a sentiment lexicon for the target language S', a word t to propagate, the score ξ of t, a diminishing factor δ, the current iteration k, and the maximum number of iterations K

output: A sentiment lexicon S' containing all propagated words with their computed sentiment scores

```
if k ≤ K then
    reachedIn = getSteps(t); // ∞ for new t
    if reachedIn > k then
        synsetsWithWord = getSynsets(synsets, t);
        foreach synset ∈ synsetsWithWord do
            pos = getPOS(synset);
            syns = getSynonyms(synset);
            foreach syn ∈ syns do
                lemma = getLemma(syn);
                sense = getWordSense(syn);
                S' = {S', {lemma, sense, pos, ξ}};
            end
            rels = getRelations(synset);
            foreach r ∈ rels do
                τ = 1;
                if r == antonym then
                    τ = −1;
                end
                rSyns = getSynonyms(r);
                foreach rw ∈ rSyns do
                    ξ' = ξτδ;
                    k' = k + 1;
                    θ = {synsets, S', rw, ξ', δ, k', K};
                    S' = propWord(θ);
                end
            end
        end
        setSteps(t, k);
    end
end
return S';
```

sentiment classification process relies on a semantic lexicon and a sentiment lexicon. We link English word senses to WordNet [55], whereas we retrieve the associated sentiment scores from SentiWordNet 3.0 [33]. On a widely used data set of 1,000 positive and 1,000 negative English movie reviews [39], our implementation has an overall sentiment classification accuracy and macro-level F_1-score of approximately 60% [34].

The implementation of our sentiment classification method for Dutch text is similar to the implementation for English text, even though it utilizes different language-specific components. For POS tagging in Dutch, we use a SharpNLP [62] POS tagger. Lemmatization is performed by the Tadpole [64] lemmatizer. Word sense disambiguation is done by applying our own implementation of the Lesk-based algorithm implemented in WordNet.Net [63]. The Dutch sentiment classifier relies on DutchWordNet (Cornetto) [57], a large semantic lexical resource for Dutch, which is used for word sense disambiguation as well as for sentiment lexicon creation using one of our considered methods other than our machine translation baseline.

We consider three main sentiment analysis approaches. In our machine translation (MT) baseline, first, we automatically translate the Dutch texts from our considered corpus into English by using the Google Translate service [65]. Then, we classify the sentiment conveyed by the translated documents by means of our sentiment classification approach for English documents.

Our first alternative to this machine translation baseline is a cross-lingual sentiment score mapping method (SMap), in which we first map the sentiment associated with all WordNet synsets from SentiWordNet 3.0 to all equivalent synsets in DutchWordNet (Cornetto). We subsequently classify the sentiment conveyed by the Dutch documents in our corpus by means of our sentiment classification approach for Dutch text, while utilizing the Dutch sentiment lexicon thus constructed.

As a second alternative to the machine translation baseline approach, we use the SProp method in order to propagate the sentiment of a set of seed words through DutchWordNet (Cornetto) and subsequently classify the conveyed sentiment by using the constructed sentiment lexicon in our sentiment classification method for Dutch documents. We assess SProp with three distinct seed sets, containing positive words (with a sentiment score of 1) and negative words (with a sentiment score of −1).

For each of these seed sets, sentiment scores are propagated by traversing the holonym, hyperonym, and hyponym relations between synsets in Dutch-WordNet (Cornetto), with a maximum number of iterations K of 8 and a diminishing factor δ of 0.9, as an initial optimization of these parameters by means of a hill-climbing procedure on one fold of our data indicated that these settings were most promising.

Each of our seed sets, detailed in Table 1, has been manually constructed by our three human annotators, all of whom are native Dutch speakers. The human annotators have combined their knowledge of the Dutch language with the most positive and negative synsets in SentiWordNet in order to construct seed sets for a Dutch sentiment lexicon. The first set contains ten positive and ten negative Dutch words. The second set is an expansion of the first set, such that it contains 26 positive and 17 negative Dutch words. Another, final expansion has resulted in a third seed set, containing 26 positive and 24 negative Dutch words.

4.2. Experimental Results

The performance of our methods of classifying the sentiment conveyed by Dutch documents by exploiting an existing method for sentiment classification of English documents is summarized in Tables 2, 3, and 4. These experimental results demonstrate that some of our approaches work better than others for performing sentiment analysis of documents in another language than the reference language. Several observations can be made from Tables 2, 3, and 4.

In general, all approaches exhibit a rather balanced performance, as they seem to perform equally well when classifying the sentiment of positive and negative documents. Additionally, when exploiting our existing sentiment analysis framework for English texts by means of our considered approaches, the best achievable performance of our framework on Dutch documents is rather comparable to the performance of the existing framework on English documents.

As reported in our previous work [34], the existing English sentiment analysis approach can obtain an overall accuracy and macro-level F_1-score of up to about 60% on a widely used collection of English movie reviews [39]. The machine translation (MT) baseline yields a sentiment classification performance on Dutch documents that is inferior to the reported performance on English documents.

Seed word	Score	Set 1	Set 2	Set 3
Mooi	1	Yes	Yes	Yes
Schoon	1	Yes	Yes	Yes
Aanbiddelijk	1	Yes	Yes	Yes
Duidelijk	1	Yes	Yes	Yes
Elegant	1	Yes	Yes	Yes
Beter	1	Yes	Yes	Yes
Glimmend	1	Yes	Yes	Yes
Perfect	1	Yes	Yes	Yes
Energiek	1	Yes	Yes	Yes
Trots	1	Yes	Yes	Yes
Super	1	No	Yes	Yes
Schitterend	1	No	Yes	Yes
Hart	1	No	Yes	Yes
Amicaal	1	No	Yes	Yes
Gezelligheid	1	No	Yes	Yes
Goed	1	No	Yes	Yes
Aanbidden	1	No	Yes	Yes
Plezier	1	No	Yes	Yes
Aangenaam	1	No	Yes	Yes
Uitmuntend	1	No	Yes	Yes
Beeldig	1	No	Yes	Yes
Positief	1	No	Yes	Yes
Veilig	1	No	Yes	Yes
Vrijheid	1	No	Yes	Yes
Vakantie	1	No	Yes	Yes
Ontspanning	1	No	Yes	Yes
Klote	-1	Yes	Yes	Yes
Boos	-1	Yes	Yes	Yes
Arrogant	-1	Yes	Yes	Yes
Bewolkt	-1	Yes	Yes	Yes
Verstoord	-1	Yes	Yes	Yes
Onmogelijk	-1	Yes	Yes	Yes
Haat	-1	Yes	Yes	Yes
Twijfelen	-1	Yes	Yes	Yes
Verafschuwen	-1	Yes	Yes	Yes
Imbeciel	-1	Yes	Yes	Yes
Mongool	-1	No	Yes	Yes
Tering	-1	No	Yes	Yes
Wantrouwig	-1	No	Yes	Yes
Verward	-1	No	Yes	Yes
Gedachteloos	-1	No	Yes	Yes
Berucht	-1	No	Yes	Yes
Jammer	-1	No	Yes	Yes
Treurig	-1	No	No	Yes
Onheilspellend	-1	No	No	Yes
Griezelig	-1	No	No	Yes
Schelden	-1	No	No	Yes
Irriteren	-1	No	No	Yes
Vervelen	-1	No	No	Yes
Negatief	-1	No	No	Yes

Table 1: Seed sets of sentiment-carrying words.

	Positive			Negative			Overall	
Method	Precision	Recall	F_1	Precision	Recall	F_1	Accuracy	F_1
MT	0.416	0.385	0.400	0.428	0.460	0.443	0.423	0.422
SMap	0.547	0.500	0.523	0.540	0.587	0.562	0.543	0.542
SProp 1	0.428	0.397	0.412	0.438	0.470	0.453	0.433	0.433
SProp 2	0.596	**0.582**	0.589	0.591	0.605	0.598	0.593	0.593
SProp 3	**0.633**	0.578	**0.605**	**0.612**	**0.665**	**0.637**	**0.622**	**0.621**

Table 2: Performance of our approaches, based on 10-fold cross-validation on our data set. The best performance is printed in bold for each performance measure.

Benchmark	MT	SMap	SProp 1	SProp 2	SProp 3
MT	0.000	0.286***	0.026	0.404***	0.471***
SMap	-0.222***	0.000	-0.202***	0.092**	0.144**
SProp 1	-0.025	0.254***	0.000	0.369***	0.435***
SProp 2	-0.288***	-0.084**	-0.270***	0.000	0.048*
SProp 3	-0.320***	-0.126**	-0.303***	-0.046*	0.000

Table 3: Relative differences of the 10-fold cross-validated overall accuracy of our approaches, benchmarked against one another on our collection of Dutch documents. Performance differences marked with * are statistically significant at $p < 0.05$, those marked with ** are significant at $p < 0.01$, and those marked with *** are significant at $p < 0.001$.

Benchmark	MT	SMap	SProp 1	SProp 2	SProp 3
MT	0.000	0.286***	0.026	0.407***	0.473***
SMap	-0.223***	0.000	-0.203***	0.094**	0.145**
SProp 1	-0.025	0.254***	0.000	0.372***	0.436***
SProp 2	-0.289***	-0.086**	-0.271***	0.000	0.047*
SProp 3	-0.321***	-0.126**	-0.303***	-0.045*	0.000

Table 4: Relative differences of the 10-fold cross-validated macro-level F_1-score of our approaches, benchmarked against one another on our collection of Dutch documents. Performance differences marked with * are statistically significant at $p < 0.05$, those marked with ** are significant at $p < 0.01$, and those marked with *** are significant at $p < 0.001$.

When using the MT method, we obtain an overall accuracy and macro-level F_1-score around a mere 47%. The SMap method yields an overall sentiment classification accuracy and macro-level F_1-score of about 54% on Dutch documents, whereas these scores amount to about 62% for SProp.

The experimental results on our corpus of Dutch documents show that our novel cross-lingual sentiment score mapping method (SMap) significantly outperforms our machine translation (MT) baseline with about 29%, caused by increased precision and recall for both positive and negative documents. Clearly, valuable information on sentiment is (partially) contained in the semantics of our source language (i.e., English), and is as such preserved when accounting for semantics by mapping the sentiment lexicon to our target language (i.e., Dutch) through relations between language-specific semantic lexicons. Accounting for semantics when propagating the sentiment of a seed set of sentiment-carrying words within a language (SProp) has even greater potential than exploiting semantics when mapping sentiment across languages. SProp significantly outperforms both MT and SMap with up to about 47% and 14%, respectively. This suggests that sentiment is not only linked to word meanings, but tends to be language-specific as well.

The machine translation approach may be thwarted by text meaning getting lost in translation. With the SMap method, noise may be introduced on word-level meanings, which apparently do not only depend on semantics, but can be language-specific as well. The SProp method is insensitive to such translation errors, as it depends on language-specific seed sets of sentiment-carrying words. The advantage of SProp does however appear to depend on the set of seed words used in the lexicon creation process. Our results suggest a sensitivity of the sentiment classification performance to the size of the seed set.

The smallest seed set, i.e., seed set 1, does not yield significant improvements over any of our methods. Conversely, a somewhat larger seed set, i.e., set 2, yields significant improvements over the MT baseline and the SProp 1 method, as well as a small, yet significant improvement over SMap. Set 3, i.e., the largest seed set, yields the largest, significant improvements over MT, SMap, SProp 1, and SProp 2. This may be explained by a larger part of the sentiment lexicon being manually annotated (i.e., the sentiment-carrying words in the seed sets), as well as by such larger initial lexicons being expanded to larger sentiment lexicons.

That the SProp 2 and SProp 3 sentiment lexicons are comparably large is clearly visible in Fig. 4, where SProp 1, SProp 2, and SProp 3 respectively cover 13%, 24%, and 29% of all terms in DutchWordNet (Cornetto), while covering 8%, 18%, and 20%, respectively, of all terms occurring in the corpus. Interestingly, the SMap lexicon yields a significantly better performance than the SProp 1 lexicon, even though the SMap lexicon only covers about 8% of the words in the corpus as well (albeit a different subset). Moreover, while covering more than two times as many of the terms occurring in the corpus, the SProp 2 lexicon significantly outperforms the SMap lexicon with only about 9%. Hence, the sentiment-carrying words in the SMap lexicon, constructed by exploiting semantic relations between languages, are comparably valuable in the analysis of the sentiment conveyed by our Dutch documents. This suggests that not only the size, but also the suitability of the seed sets for the corpus matters.

Figure 4 additionally shows that the SProp lexicons mostly cover a different part of the terms in DutchWordNet (Cornetto) than the SMap lexicon. Especially the larger SProp lexicons cover a large part of the space, in addition to the 24%, 35%, and 40% coverage of the SMap lexicon by the respective SProp 1, SProp 2, and SProp 3 lexicons. The extra coverage of the larger SProp lexicons helps improve their performance over the SMap lexicon. This confirms the importance of exploiting semantic relations within a language when constructing a sentiment lexicon.

A failure analysis has revealed that the SProp approach occasionally fails where SMap succeeds. This tends to happen when analyzing the sentiment conveyed by texts containing sentiment-carrying words that have not been assigned appropriate scores in the sentiment score propagation process.

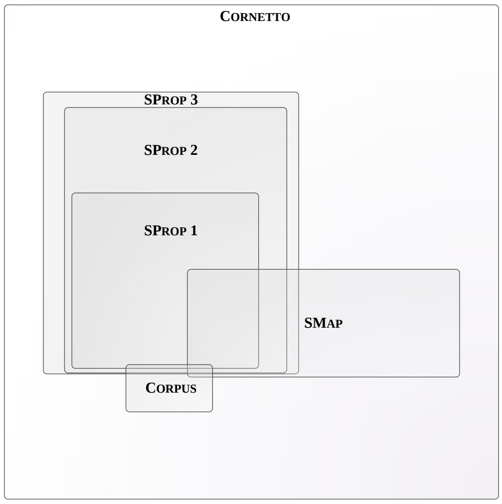

Figure 4: Coverage of the terms (i.e., unique combinations of lemmas with their parts-of-speech) in DutchWordNet (Cornetto) by those occurring in our Dutch documents, and by the *sentiment-carrying* terms in the Dutch sentiment lexicons generated by our SMap and SProp methods.

SProp may have failed to assign an appropriate sentiment score because either the associated synset was not reached by the propagation process, or the sentiment score was significantly diminished because the distance of the synset to the (possibly non-optimal) seed words was too large. Additionally, we have encountered cases in which the SMap method fails, where the SProp variants succeed. This happens when the SMap mappings do not capture the true semantics of words in Dutch, whereas the propagated SProp lexicons approximate this better.

Our failure analysis has additionally revealed that, occasionally, all of our methods fail because of misinterpreting texts. Such misinterpretations typically occur in case of negation or amplification of sentiment. Additionally, sarcasm and proverbs are interpreted literally by our current methods, as they are not covered by the resources available. Hashtags and other (misspelled) terms that are neither in our semantic lexicon nor in the constructed sentiment lexicons are another source of misinterpretations. Last, more complex structures of sentences, paragraphs, and documents are not currently taken into account. As these structures constitute the way in which sentiment-carrying words convey an author's sentiment, not accounting for these structures can cause a misinterpretation of the text in terms of its conveyed sentiment.

5. Conclusions

We have explored several methods of expanding an existing lexicon-based sentiment analysis method for a reference language, i.e., English, to another language, i.e., Dutch. Our findings suggest that, when analyzing the sentiment conveyed by texts in the target language, we cannot rely on an existing, well-performing sentiment lexicon for the reference language when simply machine-translating texts to the reference language and subsequently using the existing method for classifying the sentiment of the translated texts.

Conversely, when we map sentiment from the well-performing sentiment lexicon for our reference language to the target language by exploiting relations between language-specific semantic lexicons, we can achieve significantly better sentiment classification performance in the target language. Accounting for semantics by propagating sentiment of a seed set of sentiment-carrying words to semantically related words within the target language has even greater potential, provided that the seed set of sentiment-carrying words is sufficiently large. This indicates that sentiment is not only linked to word meanings, but tends to have a language-specific dimension as well. Thus, semantics could be exploited within a language, in addition to their use as universal link between languages when constructing sentiment lexicons in a target language.

Nevertheless, our novel sentiment mapping method, exploiting relations between language-specific semantic lexicons, has two attractive advantages over the alternative sentiment propagation method. First, in order for sentiment propagation to be truly effective, a large set of seed words in the target language is needed, whereas our sentiment mapping method does not need a seed set at all. Second, sentiment propagation is computationally more complex than our sentiment mapping method.

All in all, the key insight brought forward by our work is that semantic relations between and within languages should be carefully considered in order to exploit the full potential of sentiment analysis in real-life decision support systems that support natural language content in multiple languages. With the accuracy levels that can be obtained by our semantics-guided methods, sentiment-related information that is extracted from text in other languages than the reference language can be presented to decision makers as a rough indication of where their attention may be needed.

Our findings warrant several directions for future work. First, we could validate our findings on data in another target language. Another possible direction for future research would be to further optimize the seed sets used for the sentiment propagation process, such that they, e.g., maximize the coverage of the exploited semantic lexicon. Additionally, in future work, we could explore how to combine the sentiment propagation process with our proposed semantics-guided cross-lingual sentiment mapping approach in order to best exploit the strengths of both approaches. Last, as our findings indicate that sentiment tends to be partly language-specific, we aim to explore the comparability of sentiment scores across languages, as well as how such language-specific sentiment scores relate to an author's intended sentiment.

Acknowledgments

The authors of this paper are partially supported by the Dutch national program COMMIT. Special thanks go to Teezir (`http://www.teezir.com`) for their data and technical support.

References

[1] A. Montoyo, P. Martinez-Barco, A. Balahur, Subjectivity and Sentiment Analysis: An Overview of the Current State of the Area and Envisaged Developments, Decision Support Systems 53 (4) (2012) 675–679.

[2] C. Chang, M. Kayed, R. Girgis, K. Shaalan, A Survey of Web Information Extraction Systems, IEEE Transactions on Knowledge and Data Engineering 18 (10) (2006) 1411–1428.

[3] A. Hogenboom, F. Hogenboom, F. Frasincar, K. Schouten, O. van der Meer, Semantics-Based Information Extraction for Detecting Economic Events, Multimedia Tools and Applications 64 (1) (2013) 27–52.

[4] C. Chang, C. Hsu, S. Lui, Automatic Information Extraction from Semi-Structured Web Pages by Pattern Discovery, Decision Support Systems 35 (1) (2003) 129–147.

[5] S. Chan, Beyond Keyword and Cue-Phrase Matching: A Sentence-Based Abstraction Technique for Information Extraction, Decision Support Systems 42 (2) (2006) 759–777.

[6] L. Tari, T. Phan, J. Hakenberg, Y. Chen, S. Tran, G. Gonzalez, C. Baral, Incremental Information Extraction Using Relational Databases, IEEE Transactions on Knowledge and Data Engineering 24 (1) (2012) 86–99.

[7] W. Cui, S. Liu, L. Tan, C. Shi, Y. Song, Z. Gao, H. Qu, X. Tong, TextFlow: Towards Better Understanding of Evolving Topics in Text, IEEE Transactions on Visualization and Computer Graphics 17 (12) (2011) 2412–2421.

[8] X. Wang, X. Jin, M. Chen, K. Zhang, D. Shen, Topic Mining over Asynchronous Text Sequences, IEEE Transactions on Knowledge and Data Engineering 24 (1) (2012) 156–169.
[9] A. Balahur, J. Hermida, A. Montoyo, Detecting Implicit Experssions of Emotion in Text: A Comparative Analysis, Decision Support Systems 53 (4) (2012) 742–753.
[10] P. Lane, D. Clarke, P. Hender, On Developing Robust Models for Favourability Analysis: Model Choice, Feature Sets and Imbalanced Data, Decision Support Systems 53 (4) (2012) 712–718.
[11] A. Reyes, P. Rosso, Making Objective Decisions from Subjective Data: Detecting Irony in Customer Reviews, Decision Support Systems 53 (4) (2012) 754–760.
[12] V. Sauter, Decision Support Systems for Business Intelligence, 2nd Edition, Wiley, 2011.
[13] R. Schumaker, Y. Zhang, C. Huang, H. Chen, Evaluating Sentiment in Financial News Articles, Decision Support Systems 53 (3) (2012) 458–464.
[14] Y. Yu, W. Duan, Q. Cao, The Impact of Social and Conventional Media on Firm Equity Value: A Sentiment Analysis Approach, Decision Support Systems 55 (4) (2013) 919–926.
[15] A. Ghose, P. Ipeirotis, Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics, IEEE Transactions on Knowledge and Data Engineering 23 (10) (2011) 1498–1512.
[16] X. Yu, Y. Liu, X. Huang, A. An, Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain, IEEE Transactions on Knowledge and Data Engineering 24 (4) (2012) 720–734.
[17] S. Ludvigson, Consumer Confidence and Consumer Spending, The Journal of Economic Perspectives 18 (2) (2004) 29–50.
[18] I. Arnold, E. Vrugt, Fundamental Uncertainty and Stock Market Volatility, Applied Financial Economics 18 (17) (2008) 1425–1440.
[19] D. Baron, Competing for the Public through the News Media, Journal of Economics and Management Strategy 14 (2) (2005) 339–376.
[20] C. Holton, Identifying Disgruntled Employee Systems Fraud Risk Through Text Mining: A Simple Solution for a Multi-Billion Dollar Problem, Decision Support Systems 46 (4) (2009) 853–846.
[21] B. Jansen, M. Zhang, K. Sobel, A. Chowdury, Twitter Power: Tweets as Electronic Word of Mouth, Journal of the American Society for Information Science and Technology 60 (11) (2009) 2169–2188.
[22] R. Sprague, E. Carlson, Buidling Effective Decision Support Systems, Prentice Hall, New Jersey, USA, 1982.
[23] C. Cesarano, B. Dorr, A. Picariello, D. Reforgiato, A. Sagoff, V. Subrahmanian, OASYS: An Opinion Analysis System, in: AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs (CAAW 2006), AAAI Press, 2006, pp. 21–26.
[24] A. Devitt, K. Ahmad, Sentiment Polarity Identification in Financial News: A Cohesion-based Approach, in: 45th Annual Meeting of the Association of Computational Linguistics (ACL 2007), Association for Computational Linguistics, 2007, pp. 984–991.
[25] X. Ding, B. Lu, P. Yu, A Holistic Lexicon-Based Approach to Opinion Mining, in: 1st ACM International Conference on Web Search and Web Data Mining (WSDM 2008), Association for Computing Machinery, 2008, pp. 231–240.
[26] B. Heerschop, F. Goossen, A. Hogenboom, F. Frasincar, U. Kaymak, F. de Jong, Polarity Analysis of Texts using Discourse Structure, in: 20th ACM Conference on Information and Knowledge Management (CIKM 2011), Association for Computing Machinery, 2011, pp. 1061–1070.
[27] B. Heerschop, P. van Iterson, A. Hogenboom, F. Frasincar, U. Kaymak, Analyzing Sentiment in a Large Set of Web Data while Accounting for Negation, in: Advances in Intelligent Web Mastering - 3, 7th Atlantic Web Intelligence Conference (AWIC 2011), Vol. 86 of Advances in Intelligent and Soft Computing, Springer, 2011, pp. 195–205.
[28] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M. Stede, Lexicon-Based Methods for Sentiment Analysis, Computational Linguistics 37 (2) (2011) 267–307.
[29] A. Hogenboom, F. Boon, F. Frasincar, A Statistical Approach to Star Rating Classification of Sentiment, in: Management Intelligent Systems, 1st International Symposium on Management Intelligent Systems (IS-MiS 2012), Vol. 171 of Advances in Intelligent Systems and Computing, Springer, 2012, pp. 251–260.
[30] M. Taboada, K. Voll, J. Brooke, Extracting Sentiment as a Function of Discourse Structure and Topicality, Tech. Rep. 20, Simon Fraser University, available online, `http://www.cs.sfu.ca/research/publications/techreports/#2008` (2008).
[31] A. Hogenboom, F. Hogenboom, U. Kaymak, P. Wouters, F. de Jong, Mining Economic Sentiment using Argumentation Structures, in: Advances in Conceptual Modeling - Applications and Challenges, 7th International Workshop on Web Information Systems Modeling (WISM 2010) at the 29th International Conference on Conceptual Modeling (ER 2010), Vol. 6413 of Lecture Notes in Computer Science, Springer, 2010, pp. 200–209.
[32] A. Esuli, F. Sebastiani, SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining, in: 5th Conference on Language Resources and Evaluation (LREC 2006), European Language Resources Association, 2006, pp. 417–422.
[33] S. Baccianella, A. Esuli, F. Sebastiani, SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining, in: 7th Conference on International Language Resources and Evaluation (LREC 2010), European Language Resources Association, 2010, pp. 2200–2204.
[34] B. Heerschop, A. Hogenboom, F. Frasincar, Sentiment Lexicon Creation from Lexical Resources, in: Business Information Systems, 14th International Conference on Business Information Systems (BIS 2011), Vol. 87 of Lecture Notes in Business Information Processing, Springer, 2011, pp. 185–196.
[35] R. Mihalcea, C. Banea, J. Wiebe, Learning Multilingual Subjective Language via Cross-Lingual Projections, in: 45th Annual Meeting of the Association of Computational Linguistics (ACL 2007), Association for Computational Linguistics, 2007, pp. 976–983.
[36] B. Pang, L. Lee, Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval 2 (1) (2008) 1–135.
[37] J. Chenlo, D. Losada, Effective and Efficient Polarity Estimation in Blogs Based on Sentence-Level Evidence,

in: 20th ACM Conference on Information and Knowledge Management (CIKM 2011), Association for Computing Machinery, 2011, pp. 365–374.

[38] J. Chenlo, A. Hogenboom, D. Losada, Sentiment-Based Ranking of Blog Posts using Rhetorical Structure Theory, in: Natural Language Processing and Information Systems, Eighteenth International Conference on Applications of Natural Language to Information Systems (NLDB 2013), Vol. 7934 of Lecture Notes in Computer Science, Springer, 2013, pp. 13–24.

[39] B. Pang, L. Lee, A Sentimental Education: Sentiment Analysis using Subjectivity Summarization based on Minimum Cuts, in: 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004), Association for Computational Linguistics, 2004, pp. 271–280.

[40] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? Sentiment Classification using Machine Learning Techniques, in: Empirical Methods in Natural Language Processing (EMNLP 2002), Association for Computational Linguistics, 2002, pp. 79–86.

[41] C. Whitelaw, N. Garg, S. Argamon, Using Appraisal Groups for Sentiment Analysis, in: 14th ACM International Conference on Information and Knowledge Management (CIKM 2005), Association for Computing Machinery, 2005, pp. 625–631.

[42] G. Paltoglou, M. Thelwall, A Study of Information Retrieval Weighting Schemes for Sentiment Analysis, in: 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010), Association for Computational Linguistics, 2010, pp. 1386–1395.

[43] A. Hogenboom, P. van Iterson, B. Heerschop, F. Frasincar, U. Kaymak, Determining Negation Scope and Strength in Sentiment Analysis, in: 2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2011), IEEE, 2011, pp. 2589–2594.

[44] M. Moens, E. Boiy, A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts, Information Retrieval 12 (5) (2007) 526–558.

[45] A. Abbasi, H. Chan, A. Salem, Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums, ACM Transactions on Information Systems 26 (3) (2008) 1–34.

[46] W. Dai, G. Xue, Q. Yang, Y. Yu, Co-Clustering Based Classification for Out-of-Domain Documents, in: 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2007), Association for Computing Machinery, 2007, pp. 210–219.

[47] W. Dai, G. Xue, Q. Yang, Y. Yu, Transferring Naive Bayes Classifiers for Text Classification, in: 22nd Association for the Advancement of Articifial Intelligence Conference on Artificial Intelligence (AAAI 2007), AAAI Press, 2007, pp. 540–545.

[48] A. Gliozzo, C. Strapparava, Cross Language Text Categorization by Acquiring Multilingual Domain Models from Comparable Corpora, in: ACL Workshop on Building and Using Parallel Texts (ParaText 2005), Association for Computational Linguistics, 2005, pp. 9–16.

[49] M. Bautin, L. Vijayarenu, S. Skiena, International Sentiment Analysis for News and Blogs, in: 2nd International Conference on Weblogs and Social Media (ICWSM 2008), AAAI Press, 2008, pp. 19–26.

[50] X. Wan, Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis, in: 13th Conference on Empirical Methods in Natural Language Processing (EMNLP 2008), Association for Computational Linguistics, 2008, pp. 553–561.

[51] V. Jijkoun, K. Hofmann, Generating a Non-English Subjectivity Lexicon: Relations that Matter, in: 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2009), Association for Computational Linguistics, 2009, pp. 398–405.

[52] C. Banea, R. Mihalcea, J. Wiebe, S. Hassan, Multilingual Subjectivity Analysis Using Machine Translation, in: 13th Conference on Empirical Methods in Natural Language Processing (EMNLP 2008), Association for Computational Linguistics, 2008, pp. 127–135.

[53] C. Banea, R. Mihalcea, J. Wiebe, Multilingual Subjectivity: Are More Languages Better?, in: 23rd International Conference on Computational Linguistics (COLING 2010), Association for Computational Linguistics, 2010, pp. 28–36.

[54] Z. Lin, S. Tan, X. Cheng, A Fast and Accurate Method for Bilingual Opinion Lexicon Extraction, in: 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI 2012), IEEE, 2012, pp. 50–57.

[55] C. Fellbaum, WordNet: An Electronic Lexical Database, MIT Press, 1998.

[56] P. Vossen, EuroWordNet: A Multilingual Database for Information Retrieval, in: 3rd DELOS Workshop on Cross-Language Information Retrieval (DELOS 1997), European Research Consortium for Informatics and Mathematics, 1997, pp. 5–7.

[57] P. Vossen, K. Hofmann, M. de Rijke, E. Tjong, K. Sang, K. Deschacht, The Cornetto Database: Architecture and User-Scenarios, in: 7th Dutch-Belgian Information Retrieval Workshop (DIR 2007), Acco, 2007, pp. 89–96.

[58] M. Hu, B. Liu, Mining and Summarizing Customer Reviews, in: 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2004), Association for Computing Machinery, 2004, pp. 168–177.

[59] J. Kamps, M. Marx, R. Mokken, M. de Rijke, Using WordNet to Measure Semantic Orientations of Adjectives, in: 3rd Conference on Language Resources and Evaluation (LREC 2004), European Language Resources Association, 2004, pp. 1115–1118.

[60] K. Lerman, S. Blair-Goldensohn, R. McDonald, Sentiment Summarization: Evaluating and Learning User Preferences, in: 12th Conference of the European Chapter of the ACL (EACL 2009), Association for Computational Linguistics, 2009, pp. 514–522.

[61] T. Dao, T. Simpson, Measuring Similarity between Sentences, Tech. rep., WordNet.Net, available online, `http://wordnetdotnet.googlecode.com/svn/trunk/Projects/Thanh/Paper/WordNetDotNet_Semantic_Similarity.pdf` (2005).

[62] SharpNLP, SharpNLP, available online, `http://sharpnlp.codeplex.com/` (2006).

[63] T. Simpson, M. Crowe, WordNet.Net, available online, `http://opensource.ebswift.com/WordNet.Net` (2005).

[64] A. van den Bosch, G. Busser, W. Daelemans, S. Canisius, An Efficient Memory-Based Morphosyntactic Tagger and Parser for Dutch, in: Selected Papers of the 17th Computational Linguistics in the Netherlands Meeting (CLIN 2007), 1997, pp. 93–114.

[65] Google, Google Translate, available online, `http://translate.google.com` (2007).