

# Extracting and Grounding Context-Aware Sentiment Lexicons

Albert Weichselbraun, HTW Chur  
Stefan Gindl, Arno Scharl, MODUL University Vienna

**Abstract**—Web intelligence applications track online sources with economic relevance such as customer reviews, news articles and social media postings. Automated sentiment analysis based on lexical methods or machine learning identifies the polarity of opinions expressed in these sources to assess how stakeholders perceive a topic. This paper introduces a hybrid approach that combines the throughput of lexical analysis with the flexibility of machine learning to resolve ambiguity and consider the context of sentiment terms. The context-aware method identifies ambiguous terms that vary in polarity depending on the context and stores them in contextualized sentiment lexicons. In conjunction with semantic knowledge bases, these lexicons help ground ambiguous sentiment terms to concepts that correspond to their polarity. This grounding paves the way for interlinking, extending, or even replacing contextualized sentiment lexicons with semantic knowledge bases. An extensive evaluation applies the method to user reviews across three domains (movies, products and hotels).

**Index Terms**—H.3.1 Content Analysis and Indexing, H.3.1.d Linguistic Processing, H.2.8.f Knowledge Management Applications, H.2.8.l Text Mining

## I. INTRODUCTION

Web intelligence applications such as the *Media Watch on Climate Change* [1] shown in Figure 1 automatically elicit opinions from large text corpora, which are often extracted from online sources. News and social media in particular have turned into a crucial resource for corporate decision making. With the growing economic potential comes an academic interest to improve the accuracy of the underlying algorithms, and to integrate annotated document archives with emerging semantic knowledge bases. Sentiment analysis attracts the attention of diverse research areas including Natural Language Processing, Machine Learning and Computational Linguistics. Although a lot of effort has been put into improving of sentiment analysis, the accuracy of state-of-the-art systems still falls short of the human ability to identify opinions and infer the author’s motivation and agenda.

The limited ability of automated systems to resolve ambiguities and to process context information remains a major barrier. Context-aware sentiment analysis tackles the problem of ambiguity by attempting to determine the superordinate concept of the sentiment term in a given context. Straight-forward for humans with ample domain experience, this represents a Herculean task for an automated system. Common sense and domain-specific knowledge as well as the ability to identify relations between semantic concepts [2] are crucial to successfully solving this problem.

Building upon the technologies of the webLyzard Web intelligence platform (www.weblyzard.com), the method presented in this paper processes domain-specific corpora to learn context probabilities for the disambiguation of ambiguous sentiment terms. It extracts features from these learned context terms applicable across domains, overcoming the drawbacks of many Machine Learning approaches that are limited to their training domain. This results in extended and contextualized sentiment lexicons, which are further enriched and validated by a graph-based method that uses background knowledge from semantic knowledge bases. Third-party resources such as WordNet, ConceptNet, SenticNet, DBPedia and Freebase allow grounding ambiguous terms to the corresponding concepts. This paves the way to interlinking contextualized sentiment lexicons with semantic background knowledge and, if required, replacing such static lexicons with evolving structured data sources. Concept grounding serves as an enabler for researchers and practitioners to apply more sophisticated methods that use background knowledge. Approaches based on machine learning and lexical analysis alike can benefit from this capability to incorporate context information, for example when using constraints and identified relations between concepts to further improve the accuracy of results.

## II. RELATED WORK

Many approaches to sentiment analysis rely on sentiment lexicons, enumerative lists of sentiment terms with indicators of their sentiment charges. Popular examples include: General Inquirer [3], Subjectivity Lexicon (as used in [4]), and Subjectivity Sense Annotations [5]. SentiWordNet [6] extends the WordNet lexical database with polarity information. Balamurali et al. [7] show how to use WordNet synsets as concept features for supervised classification. While utilizing sentiment lexicons for unsupervised algorithms or feature collections for supervised classifiers delivers acceptable results, this approach fails to exploit context and complex language characteristics.

Domain knowledge plays a key role, since the linguistic context of a sentiment term often impacts its sentiment charge. Early work on sentiment detection used syntactic relations to identify new sentiment terms, which can be considered an early form of context exploitation [8]. Sentiment is often expressed in a subtle manner, which makes it difficult to identify when processing sentences or paragraphs in isolation. Context thus remains an essential ingredient to further improve sentiment analysis. Lau et al. [9] support this view by confirming that inferential language models outperform conventional models without context processing capabilities.

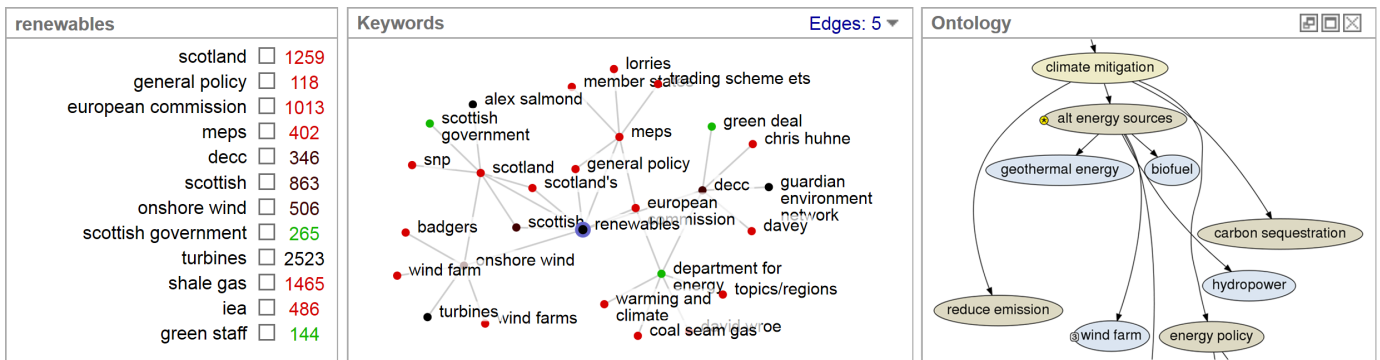


Fig. 1. Partial screenshot of the Media Watch on Climate Change ([www.ecoresearch.net/climate](http://www.ecoresearch.net/climate)); left: computed associations with "renewables" based on news media coverage between November 2011 and October 2012; center: interactive visualization of these associations, color-coded by sentiment; right: synchronized sub-graph of a climate change ontology.

Wilson et al. [4] evaluate sentence- and discourse-based valence shifters. They examine 28 syntactical and linguistic features in a machine learning approach and use these features to train four different algorithms. Larger corpora yield significantly better results than the baseline.

While the relevance of context in sentiment detection is well-established [4], research on flexible disambiguation strategies for sentiment lexicons is comparatively new. Ding et al. [10] propose rule-based context invocation to transfer polarity in compound sentences, and to surmount sentence borders based on the assumption that adjacent sentences express similar sentiment. Linguistic patterns such as " $\langle \text{object}_i \rangle$  is a little too  $\langle \text{attribute}_j \rangle$ " (e.g. "The price is a little too high") indicate a certain polarity expectation [11]. In the mentioned example, the term "price" would obtain a negative expectation.

Lu et al. [12] present an automated approach for creating context-aware sentiment lexicons based on existing lexicons and tagged consumer reviews. This lexicon contains pairs of sentiment terms and different "aspect" terms. The same sentiment term might differ in polarity when co-occurring with a particular aspect term. Gindl et al. [13] build upon this approach and present a more flexible method where pairs of sentiment terms and context terms do not receive a fixed polarity. Depending on the set of context terms contained in a document, the system calculates an overall polarity for the sentiment term. The work presented in this paper uses a refined version of this approach and extends it with capabilities to integrate third-party knowledge bases, and to extract concepts from these resources.

### III. METHOD

Sentiment lexicons contain known sentiment terms ( $t_i$ ) and their respective sentiment value ( $s(t_i)$ ). The ratio of positive and negative terms found in a document is a common indicator of overall polarity. Accurate and computationally inexpensive, it is often used as a feature for classifiers. Its accuracy can further be improved by considering linguistic features such as negations and intensifiers. The underlying assumption of stable sentiment values, however, might not hold in real-world applications. The term "perfect" is intuitively positive, for example, but switches polarity in the context of "a perfect mess". "Comedy", another positive term in its generic interpretation,

refers to a negative fact when somebody is describing a current political situation.

Disambiguation and contextualization help sentiment analysis algorithms to accurately process ambiguous sentiment terms. We identify ambiguous terms using their distribution in a labeled document corpus. Balanced occurrences in both positive and negative documents are an indicator for ambiguity. Collecting contextual data, i.e. the frequency of co-occurring terms, serves to predict the polarity in an unlabeled document. It creates a knowledge base with information on whether the term under consideration carries a sentiment value when co-occurring with certain concepts.

#### A. Context-Aware Sentiment Analysis

Figure 2 summarizes the method to create and apply context-aware sentiment lexicons. It detects ambiguous sentiment terms (Section III-A1), collects context terms for each ambiguous term (Section III-A2), and then uses the context terms to refine the sentiment analysis process (Section III-A3).

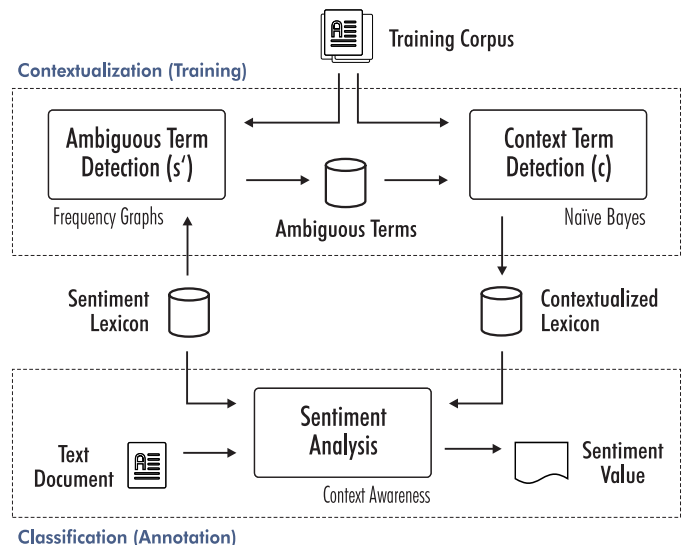


Fig. 2. Creation of contextualized sentiment lexicons.

1) *Detection of Ambiguous Terms*: The system identifies ambiguous terms using tagged training corpora. Such corpora can be created manually by reading documents and labeling them as either positive or negative. Since this time-consuming task is not feasible for large corpora, we have compiled pre-labeled corpora from online reviews (see Section IV). Such reviews are already labeled by their authors, eliminating the need for manual pre-processing. Based on these pre-labeled corpora, the system determines the distribution of each term in the sentiment lexicon. Two statistical parameters help assess the ambiguity of a term (non-ambiguous terms are used “as is” since they do not benefit from contextualization). A term is considered ambiguous if:

- its observed sentiment values show a high standard deviation ( $\sigma_{s(t_i)}$ )

$$\sigma_{s(t_i)} \geq v \quad (1)$$

- the deviations from its average sentiment value ( $\mu_{s(t_i)}$ ) lead to considerably different polarities

$$\mu_{s(t_i)} + \sigma_{s(t_i)} \geq w \quad (2)$$

$$\mu_{s(t_i)} - \sigma_{s(t_i)} \leq -w \quad (3)$$

Ambiguity assumes a certain amount of term occurrences in both polarity classes (Equation 1). Moreover, the deviation serves as a criterion for filtering neutral terms (Equation 2 and 3). Previous experiments [13] suggest threshold values of  $v = 0.75$  and  $w = 0.25$ .

2) *Collection of Context Terms*: For each identified ambiguous term, the system collects context terms and stores them in a contextualized sentiment lexicon. The number of co-occurring context terms in positive and negative documents serves as an indicator for the ambiguous terms’ positive or negative polarity. The system considers all terms, independent of their part of speech and independent of whether they represent a named entity or not. Statistical refinement removes irrelevant terms, using only context terms with the strongest probabilities for a positive or negative context. A Naïve Bayes technique (Equation 4 and 5) then estimates the polarity of an ambiguous term based on the probabilities of collected context terms  $\{c_1, \dots, c_n\}$ .

$$\mathbf{c} = \{c_1, \dots, c_n\} \quad (4)$$

$$p_t(C_+|\mathbf{c}) = \frac{p_t(C_+) \cdot \prod_{i=1}^n p_t(c_i|C_+)}{\prod_{i=1}^n p_t(c_i)} \quad (5)$$

3) *Sentiment Analysis*: Context-aware sentiment analysis combines polarity values for unambiguous and ambiguous terms, detects negation, and determines the sum of all sentiment values as the overall polarity of the document (Figure 2).

$$s_{total} = \sum_{t_i \in doc} n(t_{i-1})[s(t_i) + s'(t_i|C)] \quad \text{with (6)}$$

$$n(t_{i-1}) = \begin{cases} -1.0 & \text{if } t_{i-1} \text{ indicates a negation} \\ +1.0 & \text{otherwise.} \end{cases} \quad (7)$$

The function  $s(t_i)$  considers the contextualized sentiment lexicon and returns a term’s sentiment score. This value

becomes zero if the term is either ambiguous or not contained in the sentiment lexicon. Sentiment terms either occur in the sentiment lexicon  $s(t_i) \neq 0$  or in the contextualized sentiment lexicon  $s'(t_i) \neq 0$ . The function  $n(t_{i-1})$  detects negations and adjusts the sentiment score accordingly. By disregarding the contextualized sentiment lexicon, the same function can be used for calculating the baseline.

## B. Contextualized Lexicon Extension

Machine learning approaches tend to be corpus-specific, which can be a limiting factor when building generic opinion mining and decision support applications. Models trained on one corpus (e.g. movie reviews) might not perform as well on a corpus of a different domain (e.g. digital SLR cameras). Therefore, a specific tagged corpus is necessary for each new domain. In the case of movie and product reviews, such corpora are easy to assemble when crawled from the Web. The contextualization approach presented in this paper successfully handles more challenging domains such as climate change and political communication, where pre-tagged corpora are sparse or not available at all. The system creates a generic contextualized sentiment lexicon to be used across domains. This generic resource represents a refined lexicon merged from the contextualized lexicons of multiple corpora, distinguishing three types of context terms:

- 1) **Helpful terms** (included): Context terms classified correctly by Naïve Bayes, but not by the baseline.
- 2) **Neutral terms** (included): Context terms where both approaches yield correct result.
- 3) **Harmful terms** (disregarded): Context terms classified correctly by the baseline, but not by Naïve Bayes.

The evaluation presented in Section IV confirms that this approach yields contextualized cross-domain lexicons, which are straightforward to integrate in a wide range of opinion mining and decision support applications.

## C. Concept Analysis

Concept grounding provides a clear distinction between the concepts used in a positive and negative context, and an anchor point for integrating concepts from semantic data sources such as DBpedia, Freebase and ConceptNet. This allows interlinking, extending or even replacing the contextualized sentiment lexicon with knowledge derived from such structured resources.

Selecting the proper concept for a term is a non-trivial task. The system has to map the semantic context to external metadata. This mapping is performed by calculating the similarity between the semantic context and potential candidate concepts from third-party sources - e.g., the cosine similarity between candidate concepts from WordNet and the co-occurring terms stored in the contextualized sentiment lexicon. The sparseness and shortness of the available WordNet glosses, however, remains a challenge. Other knowledge bases such as ConceptNet or DBpedia provide even less textual information. Methods exploiting the graph structure of such repositories are an obvious way to tackle the problem. The

TABLE I  
10-FOLD CROSS VALIDATION ON THREE CORPORA AND COMPARISON OF GENERIC VERSUS DOMAIN-SPECIFIC LEXICONS

Corpus		Baseline (%)			Naïve Bayes (%)			Significance (.05)			F-Measure				Significance (.05)		
		R	P	F <sub>1</sub>	R	P	F <sub>1</sub>	p <sub>R</sub>	p <sub>P</sub>	p <sub>F</sub>	F <sub>Gen</sub>	F <sub>H</sub>	F <sub>P</sub>	F <sub>M</sub>	p <sub>R</sub>	p <sub>P</sub>	p <sub>F<sub>1</sub></sub>
Hotels	+	96	60	74	95	70	81	↓	↑	↑	81	81	74	55	↑	.	.
	-	34	90	49	55	94	69	↑	.	↑	72	69	61	71	↑	.	.
Products	+	80	64	71	78	73	75	↓	↑	.	72	69	75	43	.	.	.
	-	53	74	62	66	80	71	↑	.	↑	69	69	71	70	.	.	.
Movies	+	69	63	66	58	89	70	↓	↑	↑	75	61	66	70	↑	↑	↑
	-	60	66	63	93	69	79	↑	↑	↑	73	68	66	79	↓	↑	↓

presented work uses an approach inspired by Navigli and Lapata [14] for concept disambiguation. The system creates a graph from the knowledge base, where nodes represent concepts and edges the relations between those concepts. This graph yields connectivity measures between potential concepts (e.g., WordNet senses) and the senses of its context terms that are translated into corresponding similarity measures:

$$sim(s_1, s_2) = \frac{1}{dist_{min}(s_1, s_2) + 1} \quad (8)$$

The process outlined below identifies the WordNet sense of the ambiguous sentiment term based on its context terms (i.e., all other terms in the sentence or paragraph not contained in a stopword list) by obtaining a list of WordNet senses for the ambiguous term (line 3), and computing the similarity ( $sim[sense]$ ) between each sense and the context terms. After WordNet senses for each context term (line 6) have been retrieved, the algorithm (lines 7-13) determines the similarity ( $maxContextSim$ ) between the current sense ( $sense$ ) and the best matching sense of the current context term ( $contextTerm$ ). Aggregating the similarity values between the ambiguous term and the closest senses of all context terms then yields  $sim[sense]$ . The system chooses the WordNet sense with the strongest connection to the context terms, i.e., the  $sense$  maximizing  $sim[sense]$ .

```

1: ambiguousTerm ← String
2: contextTerms ← [posContextTerm1, posContextTerm2, ...] or
   [negContextTerm1, negContextTerm2, ...]
3: senses ← getWordnetSenses(ambiguousTerm)
4: for all sense in senses do
5:   for all contextTerm in contextTerms do
6:     contextTermSenses ← getWordnetSenses(contextTerm)
7:     maxContextSim ← 0
8:     for all contextTermSense in contextTermSenses do
9:       similarities ← getSim(sense, contextTermSense)
10:      if sim > maxContextSim then
11:        maxContextSim ← sim
12:      end if
13:    end for
14:    sim[sense] ← sim[sense] + maxContextSim
15:  end for
16: end for

```

#### IV. EVALUATION

The evaluation draws upon 2 500 product reviews from *Amazon.com*, 1 800 hotel reviews from *TripAdvisor.com* and the movie review corpus of Pang and Lee [15]. The ratings from these sites range from 1-5 stars (or circles). Ratings

with fewer (more) than three stars were considered negative (positive). Each of the three corpora was compiled with an equal number of positive and negative reviews. The corpora will be referred to as *Products*, *Hotels*, and *Movies* in the remainder of this paper.

The goal of the evaluation is to compare domain-specific with generic (i.e., training and testing across domains) contextualization, assess the performance of the context term selection process, and illustrate the plausibility of the extracted concepts. The experiments helped to verify the three hypotheses outlined in the following sections.

DOMAIN-SPECIFIC CONTEXTUALIZATION (*Hypothesis 1*). *Context knowledge improves lexicon-based sentiment analysis. A contextualized sentiment lexicon trained on corpus A with the proposed Naïve Bayes method delivers superior results on corpus A as compared to the baseline lexicon.*

Table I shows that the contextualized lexicons outperform the original lexicons. The increase in F-measure (↑) is significant across corpora and polarities, except a non-significant improvement (·) in the case of positive product reviews.

In three cases, the detection of positive reviews in the *Products* and *Movies* corpora, the contextualized lexicon reduced recall (indicated by ↓), but only one of these reductions was significant at the 0.05 level (based on Wilcoxon’s rank sum test). This result is not surprising, since improvements in precision often lead to a slight decay of recall and vice versa. The F-measure as a hybrid measure combines precision and recall to judge the overall performance of the algorithm.

GENERIC CONTEXTUALIZATION (*Hypothesis 2*). *Identifying generic context terms improves the applicability to other domains. A contextualized lexicon with generic context terms from corpora A<sub>1</sub> to A<sub>n</sub> yields better results on corpora A<sub>1</sub> to A<sub>n</sub> than a contextualized lexicon trained on a single corpus.*

Merging the context terms of contextualized sentiment lexicons originating from three different corpora provides generic context terms that are useful across domains. We created a generic contextualized sentiment lexicon as outlined in Subsection III-B and evaluated it on three test corpora.

Domain-specific lexicons typically deliver superior results in their training domain, but do not perform well when applied across domains. Table I compares the F-measures of the generic lexicon with the results obtained using domain-specific lexicons, using Wilcoxon’s rank sum test to compute significance. Arrows indicate a statistically significant increase (↑) or decrease (↓). The results based on the generic lexicon do not differ significantly - the *Movies* corpus being the only

TABLE II  
CONTEXTUALIZATION EXAMPLES

Ambiguous Term	Sentiment Value	Example Sentence
<b>busy</b>	1	The hotel is located on a <i>busy road</i> .
<b>complaint</b>	-1	My <i>only complaint</i> would be the service.
<b>cool</b>	-1	Our room felt like a <i>really cool</i> european apartment with a rooftop terrace.
<b>expensive</b>	-1	The room was one of the more <b>expensive</b> hotels in Vienna but still <i>excellent</i> .
<b>quality</b>	1	<i>Poor quality</i> copies with one edge always dark.
<b>better</b>	1	Let's <i>hope</i> they work <b>better</b> .
<b>cost</b>	-1	Toner <b>cost</b> is way <i>behind</i> competitors.

exception, where an improvement for positive reviews is offset by a decrease in F-measure for negative reviews.

We conducted similar experiments with the standard Naïve Bayes classifier of NLTK, the Natural Language Toolkit (www.nltk.org). The evaluation confirmed that cross-domain sentiment analysis remains a challenging task. The classifier performed very well on the domain it had been trained on, but achieved only poor results when applied to other domains.

CONCEPT ANALYSIS (*Hypothesis 3*). *The integration of context terms extracted from the contextualized sentiment lexicon with background knowledge from WordNet separates ambiguous sentiment terms into positive and negative concepts and thereby helps to explain inherent ambiguities.*

Table II exemplifies the identification of context terms during the sentiment analysis process. The left and middle columns contain the ambiguous term and its sentiment value, as stored in the sentiment lexicon. The right column shows a sentence, where the context term (in italic) inverted the sentiment value of the sentiment term (in bold). “busy”, for example, is a sentiment term with a positive value in the sentiment lexicon. Contextualization changes its polarity when used in the context of “busy roads”. The term “cool” (negative in the initial sentiment lexicon) becomes positive in conjunction with “really”. Similarly, the positive term “quality” changes its value when used in the context of “poor quality”. One could argue that these improvements could have been achieved with different techniques such as n-grams as well. N-grams would address the case of “poor quality”, but fail in sentences such as “The quality of device X is poor”. Contextualization goes beyond individual improvements and optimizations. It is an effective method for addressing a broad variety of natural language rules and exceptions by incorporating external knowledge not only from various domain corpora, but also from structured semantic archives.

After adding context knowledge to the original sentiment lexicon, we applied concept analysis to identify meaningful concepts based on the extracted context terms. The concept-aware sentiment component logs both successful (e.g. “*Disambiguating charged - Synset(charge.v.23), Synset(charge.v.08)*”) and unsuccessful (e.g. “*Cannot Disambiguate admirer - both concepts refer to Synset(admirer.n.02)*”) disambiguation attempts. Table III summarizes the results and sheds light on the quality of the disambiguation process. Camera or hotel reviews tend to use a more consistent vocabulary than movie reviews, which cover many different aspects including the storyline and

TABLE III  
WORDNET CONCEPTS EXTRACTED FOR AMBIGUOUS TERMS

Amazon	2287 of 5259 (43%) sentiment terms disambiguated	
<b>gone</b>	<i>go.v.10</i> have the quality of being; be	<i>fail.v.04</i> render inoperable/ineffective blow out, malfunction, crash, die, misfire, break, change
<b>overcome</b>	<i>get_the_better_of.v.01</i> come out better in a competition, race, or conflict demolish, beat, rout, upset, lurch, down, conquer, nose, wallop, survive, overrun	<i>overwhelm.v.01</i> call forth (emotions, feelings, and responses) benight, knock out, devastate, seize, lock, stagger, arouse, kill
<b>charged</b>	<i>charge.v.23</i> energize a battery by passing a current through it charge, change	<i>charge.v.08</i> saturate change, freight, burden
TripAdvisor	1881 of 3735 (43%) sentiment terms disambiguated	
<b>calm</b>	<i>calm.v.01</i> give emotional strength pacify, compose, reassure, lull, comfort	<i>composure.n.01</i> your usual mood aplomb, disposition, repose
<b>charming</b>	<i>capture.v.02</i> be attractive to hold, work, attract	<i>charm.v.02</i> exercise authoritative control or power over control, hex
<b>worry</b>	<i>worry.n.02</i> unpleasant emotion in anticipation of some (usually ill-defined) misfortune anxiety	<i>concern.n.04</i> a stimulus with undesirable consequences bugaboo, negative stimulus, burden, business
IMDB	10437 of 31955 (33%) sentiment terms disambiguated	
<b>jerk</b>	<i>jerk.n.02</i> change of position not entailing a change of location motion	<i>jerk.n.01</i> someone unable to adapt to their circumstances misfit, schmuck
<b>loving</b>	<i>love.v.03</i> have a great affection for romance, love	<i>sleep_together.v.01</i> kiss, embrace, fondle neck, take, fornicate
<b>dumb</b>	<i>speechless.s.01</i> temporarily incapable of speaking	<i>dense.s.04</i> slow to learn or understand; lacking intellectual acuity

its social context, the performance of the actors, and the overall quality of the production. This heterogeneity is reflected in the concept analysis performance, which performs better in domains with consistent vocabularies. A qualitative analysis of the concept grounding process and the successfully disambiguated WordNet term definitions showed that the identified concepts in Table III reflect subtle nuances in term usage and their impact on the author’s attitude towards a subject.

## V. CONCLUSION AND FUTURE WORK

This paper presents a method to improve sentiment analysis by using contextualized sentiment lexicons for disambiguating sentiment terms. A graph-based component for concept identification refines these lexicons and uses WordNet to ground ambiguous sentiment terms to concepts. This grounding process provides a clear distinction between positive and negative concepts and paves the way for incorporating semantic databases into the sentiment analysis process. Analogous to distinguishing context terms based on part-of-speech information and co-occurrence patterns, the adaptive computation of sentiment

scores based on confirmed associations with WordNet concepts (synsets) positively impacts the accuracy of results as well as the generic applicability of lexicons across domains.

The presented approach and data representation for extending existing sentiment lexicons with data from structural sources (or even replacing lexical data with structured knowledge) is highly flexible. Context-aware sentiment analysis and concept analysis as outlined in this paper are independent from particular domains and pre-defined linguistic patterns; they analyze the frequency distribution of context terms co-occurring with an ambiguous sentiment term in a labeled corpus. Based on learned probabilities, the method predicts the polarity of an ambiguous sentiment term given the word frequencies in an unlabeled document.

Extending classic sentiment lexicons with this context information creates domain-specific lexicons, and identifies the concepts that correspond to these lexicons. We then create domain-independent lexicons by discarding terms potentially harmful for a correct disambiguation based on cross-domain evaluations. Although these lexicons cannot compete with tailored domain-dependent machine learning methods such as Naïve Bayes in their training domain, they significantly outperform these approaches across domains.

The performance of the concept grounding process depends on the complexity of the observed domain, typically yielding better results for domains that discuss a rather consistent topic such as photo cameras (about 50% of the concepts were disambiguated successfully, as compared to 33% in the case of more complex domains such as movie reviews). The method returns intuitive concept distinctions, correctly processes subtle nuances of natural languages, and helps understand semantic relations in contextualized sentiment lexicons.

Future research will focus on applying the presented method to more comprehensive semantic knowledge bases such as ConceptNet, DBpedia and Freebase, and on using the grounded concept in conjunction with these sources for refining and enriching the contextualized sentiment lexicons to further increase the achievable level of generalization.

#### ACKNOWLEDGEMENT

The research presented in this paper has been conducted as part of the DIVINE Project ([www.weblyzard.com/divine](http://www.weblyzard.com/divine)), funded by the Austrian Research Promotion Agency within the FIT-IT Semantic Systems program line, and the WISDOM Project ([www.semanticlab.net/wisdom](http://www.semanticlab.net/wisdom)) funded by the Swiss Commission for Technology and Innovation.

#### REFERENCES

- [1] A. Hubmann-Haidvogel, A. Scharl, and A. Weichselbraun, "Multiple Coordinated Views for Searching and Navigating Web Content Repositories," *Information Sciences*, vol. 179, no. 12, pp. 1813–1821, 2009.
- [2] A. Weichselbraun, G. Wohlgenannt, and A. Scharl, "Refining Non-Taxonomic Relation Labels with External Structured Data to Support Ontology Learning," *Data & Knowledge Engineering*, vol. 69, no. 8, pp. 763–778, 2010.
- [3] P. J. Stone, *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, Massachusetts, U.S.A.: MIT Press, 1966.
- [4] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis," *Computational Linguistics*, vol. 35, no. 3, pp. 399–433, Sep. 2009.
- [5] J. M. Wiebe and R. Mihalcea, "Word Sense and Subjectivity," in *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics*. Sydney, Australia: ACL, 2006, pp. 1065–1072.
- [6] A. Esuli, F. Sebastiani, and V. G. Moruzzi, "SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining," in *Proceedings of 5th Conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy, 2006, pp. 417–422.
- [7] A. R. Balamurali, A. Joshi, and P. Bhattacharyya, "Harnessing WordNet Senses for Supervised Sentiment Classification," in *Proceedings of the Conference on Empirical Methods in Natural Language (EMNLP 2011)*, 2011, pp. 1081–1091.
- [8] V. Hatzivassiloglou and K. R. McKeown, "Predicting the Semantic Orientation of Adjectives," in *Proceedings of the 8th Conference on the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics. Madrid, Spain: ACL, 1997, pp. 174–181.
- [9] R. Y. Lau, C. Lai, and Y. Li, "Leveraging the Web Context for Context-sensitive Opinion Mining," in *Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2009)*. Beijing, China: IEEE, 2009, pp. 467–471.
- [10] X. Ding, B. Liu, and P. S. Yu, "A Holistic Lexicon-based Approach to Opinion Mining," in *Proceedings of the 1st International Conference on Web Search and Data Mining (WSDM 2008)*. Palo Alto, CA, USA: ACM Press, 2008, pp. 231–239.
- [11] Y. Wu, M. Wang, and P. Jin, "Disambiguating Dynamic Sentiment Ambiguous Adjectives," in *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, 2010, pp. 1191–1199.
- [12] Y. Lu, M. Castellanos, U. Dayal, and C. Zhai, "Automatic Construction of a Context-Aware Sentiment Lexicon: An Optimization Approach," in *Proceedings of the 20th International Conference on the World Wide Web (WWW 2011)*. Hyderabad, India: ACM, 2011, pp. 347–356.
- [13] S. Gindl, A. Weichselbraun, and A. Scharl, "Cross-Domain Contextualisation of Sentiment Lexicons," in *19th European Conference on Artificial Intelligence (ECAI 2010)*, ser. Frontiers in Artificial Intelligence and Applications, H. Coelho, R. Studer, and M. Wooldridge, Eds., vol. 215. Lisbon, Portugal: IOS Press, August 2010, pp. 771–776.
- [14] R. Navigli and M. Lapata, "Graph Connectivity Measures for Unsupervised Word Sense Disambiguation," in *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*, 2007, pp. 1683–1688.
- [15] B. Pang and L. 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 (ACL 2004)*. ACL, 2004, pp. 271–278.

#### THE AUTHORS

Albert Weichselbraun is a Professor of Information Science at the University of Applied Sciences Chur and Technical Director of the webLyZard technology. His research focuses on the integration of text mining and semantic technologies for ontology evolution, information diffusion, sentiment analysis and Web intelligence. Contact him at [albert.weichselbraun@htwchur.ch](mailto:albert.weichselbraun@htwchur.ch).

Stefan Gindl is a researcher and lecturer at MODUL University Vienna's Department of New Media Technology, and a founding member of the Interest Group on German Sentiment Analysis. His research focuses on sentiment analysis with an emphasis on context-awareness and incorporating common sense into existing approaches. Contact him at [stefan.gindl@modul.ac.at](mailto:stefan.gindl@modul.ac.at).

Arno Scharl heads the Department of New Media Technology at MODUL University Vienna and is the Managing Director of webLyZard technology. His research interests include media monitoring and Web intelligence, human-computer interaction, and the integration of semantic and geospatial Web technologies. Contact him at [scharl@modul.ac.at](mailto:scharl@modul.ac.at).