

# Context-sensitive Sentiment Propagation in WordNet

**Jan Kocon**  
G4.19 Research Group  
Wrocław University  
of Science and Technology  
Wrocław, Poland  
jan.kocon@pwr.edu.pl

**Arkadiusz Janz**  
G4.19 Research Group  
Wrocław University  
of Science and Technology  
Wrocław, Poland  
arkadiusz.janz@pwr.edu.pl

**Maciej Piasecki**  
G4.19 Research Group  
Wrocław University  
of Science and Technology  
Wrocław, Poland  
maciej.piasecki@pwr.edu.pl

## Abstract

In this paper we present a comprehensive overview of recent methods of the sentiment propagation in a wordnet. Next, we propose a fully automated method called Classifier-based Polarity Propagation, which utilises a very rich set of features, where most of them are based on wordnet relation types, multi-level bag-of-synsets and bag-of-polarities. We have evaluated our solution using manually annotated part of plWordNet 3.1 emo, which contains more than 83k manual sentiment annotations, covering more than 41k synsets. We have demonstrated that in comparison to existing rule-based methods using a specific narrow set of semantic relations our method has achieved statistically significant and better results starting with the same seed synsets.

## 1 Introduction

Princeton WordNet (Miller, 1995) has been expanded with sentiment annotation in several projects. However in all these approaches only a very limited part of Princeton WordNet was manually annotated, and the annotation for the remaining part was automatically extended by propagation algorithms, e.g. WordNet-Affect (Strapparava and Valitutti, 2004) or SentiWordNet (Esuli and Sebastiani, 2006), see also Sec. 3. Manual emotive annotation was done for plWordNet (Maziarz et al., 2016) (a wordnet of Polish) on several times larger scale. In the most contemporary version more than 54 000 lexical units (i.e. word senses) are described by sentiment polarity, basic emotions and fundamental human values, cf. (Zaśko-Zielińska et al., 2015). Only nouns and adjectives are annotated, but the manual annotation coverage

of these two part-of-speech categories is almost 24%. Having this large amount of metadata we started to look at methods of automated expansion of such information in a wordnet. Most of the existing solutions are based on a set of handcrafted rules for transferring the polarity along different types of wordnet relations. The proposed method does not require manually designed rules as they are discovered automatically.

## 2 Background

Lexicons are an important, inherent part of sentiment analysis and opinion mining systems. There are three general approaches to compile sentiment lexicon i.e. *corpus-based* approach: *dictionary-based* and *manual* (Liu, 2015). Manual approaches are laborious and time-consuming, so there is a great need for fast, automated methods of the construction of sentiment lexicons especially for low-resourced languages. The first built lexicons were limited only to simple word lists with positive and negative examples of words. However, the polarity of words often varies across their senses due to the semantic ambiguity. We assume that a sense-based sentiment lexicon may enable more accurate estimation of the sentiment polarity of complex phrases or sentences. One of the possible ways to construct a sense-aware sentiment lexicon is to use a wordnet (i.e. a dictionary-based approach). Approaches of this kind of generally aim at extending a small set of seed words with known polarity using lexical relations of a wordnet, e.g. hypernymy, synonymy, antonymy, etc.

Most of the existing solutions rely on a simple polarity propagation from annotated synsets (*seeds*) to their not annotated neighbours, and mostly utilise only specific subset of relations like hypernymy, hyponymy, similarity and antonymy (Maks and Vossen, 2011). These approaches do not take into account the full structure of WordNet

or even wider contexts of synsets (e.g.  $n$ -th level relations). A common approach to construct a non-English sentiment lexicon is a simple translation of SentiWordNet (Esuli and Sebastiani, 2006) polarity annotations to another language.

Simple rule-based propagation prepared for one language does not necessarily perform well for other languages, because wordnets for different languages may differ strongly, e.g. in the number of relation instances and a different semantic structure. On the other hand, corpus-based solutions require a high quality systems for word sense disambiguation. A good method for sentiment propagation should be adaptable to the structure of any wordnet with the least human effort.

### 3 Related Works

There is a vast amount of methods to construct sentiment lexicons, but most of them were evaluated only for English, on Princeton WordNet (Miller, 1995). One of the major sentiment lexicons for English – SentiWordNet – was introduced in (Esuli and Sebastiani, 2006), and in (Baccianella et al., 2010) its extended version was described. The main objective was to construct a large lexical resource with sentiment polarity of lexical meanings rather than words.

One of the approaches based on a non-English wordnet was evaluated in (Maks and Vossen, 2011). The authors compared three methods:

1. Simple polarity transfer from SentiWordNet (Esuli and Sebastiani, 2006) using translation equivalents between Princeton WordNet and Dutch WordNet (Piek Vossen and VanderVliet, 2008);
2. Automatic polarity propagation using only Dutch WordNet;
3. Combined approach using transfer method from SentiWordNet and polarity propagation over the Dutch WordNet.

The first method resulted in a general performance decrease in comparison to SentiWordNet from 62% to 58% of overall precision, recall and F-score. The second method was based on iterative label propagation with rules using lexical relations from WordNet. Factors such as seed set size, its composition and number of iterations had a great impact on propagation performance. When high-quality pre-selected seed synsets are used,

the obtained performance is significantly higher. One of the drawbacks of their approach is the simplicity of seed selection criteria. The best results were achieved using a mixed dataset derived from a large sentiment lexicon – the General Inquirer Lexicon (Stone, 1966). The performance reached 75% of F-score, precision and recall. The authors concluded that the size of a seed set is the most important factor, but the quality of the seeds also matters. Almost the same performance was achieved by combining transfer method with propagation (74%). The results may also suggest that simple transfer methods are not perfect, but combining multiple approaches with transfer methods may bring us a promising result.

Extended research on the polarity propagation for non-English wordnets was presented in (Maks et al., 2014). Authors applied the same propagation algorithm to five wordnets for different languages. The propagation method was similar to the methods used in their previous works. Words and their polarity extracted from the well-known General Inquirer Lexicon were translated with a machine translation service and manually mapped to the corresponding synsets in particular wordnets. The seed set consisting of synsets with known polarity was expanded using wordnet relations to cover the entire networks. The resulting lexicons varied significantly in their size and precision score. The conclusion was that the way the wordnets are built seems to affect propagation performance.

(Mahyoub et al., 2014) is a first attempt to build an Arabic sentiment lexicon on a basis of Arabic WordNet. Propagation procedure involves an expansion step which is expanding the sentiment lexicon by iteratively reaching concepts of the wordnet and scoring step evaluating the sentiment score of reached concepts according to their distance from the seeds. A task-based evaluation was applied. The acquired polarity scores were incorporated into features for sentiment classification task evaluated on Arabic corpora.

There were several attempts to construct a large sentiment lexicon for Polish in an automated way e.g. (Haniewicz et al., 2013; Haniewicz et al., 2014). (Haniewicz et al., 2013) attempted to build a polarity lexicon from web documents. They utilized plWordNet (yet without sentiment annotation) as a general resource to develop domain-aware polarity lexicons. A large semantic lexicon

with over 70,000 concepts from Web reviews was built where each term in this lexicon was described by a vector of sentiment values, representing the polarity of this term in various domains. plWordNet was utilised to identify semantic relations between acquired terms. To determine their polarity a supervised learning with Naive Bayes and SVM was applied. This approach was extended in (Haniewicz et al., 2014) where the semantic lexicon was expanded to 140,000 terms. To enlarge the lexicon the authors used a simple rule-based propagation with an adaptation of Random Walk algorithm.

SentiWordNet construction in its recent stages was generally based on the glosses from Princeton WordNet. (Misiaszek et al., 2013) proposed a lexicon construction method for wordnets, for which a simple transfer method could not be easily applied or external sources of knowledge such as tagged and disambiguated glosses are not available. This approach was based on relational propagation scheme with local, collective classification method, namely Iterative Classification Algorithm (ICA) for determining polarity of synsets. The training features for the classifier were obtained using only a neighbourhood of annotated synsets, consisting of nodes with known polarity. They manually annotated specific synsets in the wordnet and used them as seeds for the propagation process. However, the details of the feature extraction were not specified and there was no evaluation for their approach.

In (Kulisiewicz et al., 2015) the propagation was performed by using an adaptation of Loopy Belief Propagation (LPA) on Princeton WordNet 3.0. Three different variants of the LPA have been tested and evaluated. The evaluation was carried out in two ways. Firstly, the authors compared their results with polarity scores from SentiWordNet (*Mean Square Error*), but skipping the *Objective* class. Secondly, evaluation was done by comparison with polarity of words existing in the General Inquirer Lexicon. The resultant performance was ambiguous, and the main conclusion was that semantic relations within wordnet may not be a well correlated with the sentiment relations.

#### 4 Classifier-based Polarity Propagation

We propose a fully automated method called *Classifier-based Polarity Propagation* (henceforth CPP) with a very rich set of features. In Sec-

tion 5.1 we compare the results obtained by CPP with rule-based and relation-based method called *Seed Propagation* and its best configuration presented by Maks and Vossen (2011).

##### 4.1 Polarity Transfer from Units to Synsets

We analysed the contemporary annotation of plWordNet to see how diverse synsets are in terms of units polarity. In contrast to SentiWordNet the manual annotation in plWordNet is done on the level of lexical units (Zaśko-Zielińska et al., 2015). Available values for polarity are: *strong negative*, *weak negative*, *neutral*, *weak positive*, *strong positive*, *ambiguous*. One annotator can assign only one of these values for a single lexical unit.

Currently there are more than 83k annotations covering more than 54k lexical units and 41k synsets. About 22k of the polarity annotations are different than neutral and these annotations cover 13k lexical units and 9k synsets (22% of all synsets including annotated units). We found that 1.5k of these synsets were annotated with different polarity across their units. If we exclude neutral units, only 345 of them have varying polarity strength (e.g. synset that contains two lexical units annotated as strong positive and one annotated as weak positive). If we exclude both neutral and ambiguous annotations, there is only 41 synsets having conflicting, opposite polarity of their units (synsets that have both positive and negative units), and it is only 3.8% of all polarized synsets (synsets that do not contain any neutral units - 9164).

The acquired statistics show that synsets are strongly homogeneous in terms of the lexical units polarity, so we decided to move annotations from unit-level to the synset-level. In order to simplify the problem we decided to project these values to only three: *positive*, *negative*, *neutral*. For each annotation value we assigned the following weights: 2 for *strong* variants, 1 for weak variants, neutral and ambiguous. Then we recounted the number of annotations in each synset including assigned weights. For example, if we have a synset with a set of its lexical units like {*strong negative*, *negative*, *strong positive*, *neutral*}, we have total weight for *positive* category equal to 2, *negative* category equal to 3 (2 + 1) and *neutral* category equal to 1. We decided to assign only one polarity class to each synset – the one having the largest

Relation	Occurrences [%]
hyponymy	34.72
hypernymy	34.72
fuzzynymy	9.40
similar_to	3.20
feature_value	3.03
meronymy	1.86
holonymy	1.49
collection_meronym	1.29
collection_holonym	1.23
type	1.06
member	1.06
taxonomic_meronym	1.00
taxonomic_holonym	0.99
SUM	95.06

Table 1: Frequency (as part of the whole number of relations) of the selected relations in plWordNet.

weight. In the given example the assigned polarity will be *negative*. If we have two classes of the same weight, we apply the following rules to solve this discrepancy:

- $\{positive, neutral\} \rightarrow positive$
- $\{negative, neutral\} \rightarrow negative$
- $\{positive, negative\} \rightarrow neutral$

## 4.2 Features

We analysed the existing structure of plWordNet to select the most common relations. The results are presented in Table 1. We took a subset of relations which covers more than 95% of all relation instances in plWordNet.

Each synset is described by a set of features, where the feature value is represented as *bag-of-words* containing synsets or polarities. Each feature type is a set of 4 variables:

- *Relation* – one of the 13 relations given in Table 1.
- *Direction* – the direction of the relation, the described synset can be a *source* or *target* of the given relation.
- *Word* – there are two types of *words* in *bag-of-words* model: *synset\_ID* (any number) and *synset\_polarity* (one of the following numbers:  $-1, 0, 1$ ; it represents 3 polarity classes: *negative, neutral, positive*).

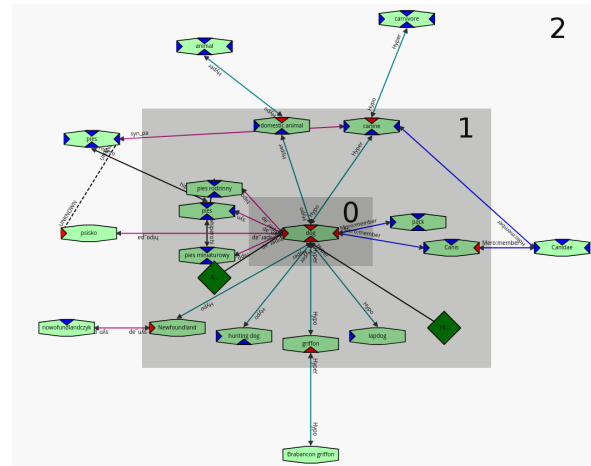


Figure 1: Example of synsets at the specific level (1 and 2) with respect to the synset at level 0.

- *Level* – the first level means synsets in *direct* relation to the described synset, the second level are synsets in direct relation with synsets from the first level, but excluding synsets from the first level. The example is presented in Figure 1.

There are 13 relations, 2 directions, 2 word types and 2 levels, which in total gives  $13 \cdot 2^3 = 104$  types of features. For example a feature of the type *hyponym\_source\_id\_level\_2* contains all IDs of synsets which are sources of all hyponym relation instances, for which the target is any synset at the 1st level (see Figure 1).

## 4.3 Classifier

Having a set of annotated synsets and 104 bags of words as features for each synset, we utilised `TfidfVectorizer` module from `scikit-learn`<sup>1</sup> Python machine learning package. This feature extraction method allows to convert a collection of elements to a matrix of TF-IDF features. Each synset belongs to one of three following classes: *positive, negative, neutral*. Transformed data is used to train a predictive model. We used Logistic Regression from `scikit-learn` package as a classifier.

## 4.4 Propagation

With a trained classifier we perform propagation for the remaining, unlabelled part of plWordNet. At the beginning we treat our seeds as a set of synsets at level-0 (see figure 1). Each next iteration is a classification of synsets at the 1st level,

<sup>1</sup><http://scikit-learn.org>

using annotated synsets from the other levels. We prepared the solution using one of the following approaches applied to each iteration:

- naive – we get the graph order of the remaining synsets to be classified,
- sorted – before each iteration we sort synsets at the 1st level by the number of relations with synsets which already have the polarity value assigned (descending order).

## 5 Experiments and Results

### 5.1 Experimental Set-up

The developed method assumes that the propagation is performed only for synsets. However, existing polarity annotations in the plWordNet refer only to lexical units, thus some pre-processing was required. First, we used a simple generalization function to assign the polarity to the synsets, depending on the polarity of their units (see Section 4.1) and projecting a 5-degree scale of polarity to a 3-degree scale. Then, to evaluate the lexicon we prepared a large graph of plWordNet, consisting of generalized synsets.

### 5.2 Evaluation Procedure

The evaluation procedure utilises full plWordNet with 43k synsets annotated with sentiment polarity (positive, negative, neutral). Annotated synsets were divided once into 10 parts and 9 parts (about 40,400 synsets) were treated as a seed set for baseline (or learning set for CPP) and the last part (about 3,600 synsets) as a test set. For each method and configuration we performed 10-fold cross-validation.

We implemented a simple rule-based seed-driven propagation method described in (Maks and Vossen, 2011) to obtain a *baseline* (henceforth BASE). Then we compared the results with CPP in two variants described in Section 4.4: naive (CPP-N) and sorted (CPP-S).

### 5.3 Results and Discussion

Table 2 presents the results obtained during experiments. We calculated precision (P), recall (R) and F-measure (F) for separate classes of polarity: negative (NEG), positive (POS) and neutral (NEU). We compared differences between two pairs: {BASE, CPP-N} and {CPP-N, CPP-S}. In Tab. 2 we highlighted results for which differences were statistically significant. We anal-

Measure	BASE	CPP-N	CPP-S
P-NEG	84.01	84.58	84.73
P-NEU	92.18	<b>93.75</b>	93.66
P-POS	69.20	<b>83.11</b>	82.95
R-NEG	68.63	<b>75.82</b>	75.90
R-NEU	95.80	<b>97.02</b>	96.97
R-POS	64.64	<b>68.41</b>	67.80
F-NEG	75.52	<b>79.91</b>	79.81
F-NEU	93.95	<b>95.34</b>	95.35
F-POS	66.77	<b>74.99</b>	74.61

Table 2: Precision (P), recall (R) and F-score (F) for separate classes of polarity. BASE results are compared to CPP-N and CPP-S. Statistically significant differences are emphasised.

ysed the statistical significance of differences using paired-differences Student’s t-test with a significance level  $\alpha = 0.05$  (Dietterich, 1998).

Naive solution (CPP-N) is significantly better than BASE in all test cases except precision for class *negative*. The order of neighbours classified in each iteration is not important in this case, because there was no significant difference between CPP-N and CPP-S variants.

## 6 Conclusions

The results prove that the proposed method performs better in almost all cases comparing to simple rule-based methods which transfer known polarity from seeds to other parts of wordnet. Surprisingly for us, the solution with sorting synsets in each iteration in descending order by the number of neighbours with known polarity did not provide any increase of propagation quality. We think that the further work should be concentrated on training the classifier after each iteration and in this scenario sorting before classifying should be beneficial.

## Acknowledgments

Work co-financed as part of the investment in the CLARIN-PL research infrastructure funded by the Polish Ministry of Science and Higher Education and in part by the National Centre for Research and Development, Poland, under grant no POIR.01.01.01-00-0472/16.

## References

- [Baccianella et al.2010] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the International Conference on Language Resources and Evaluation*, volume 10, 01.
- [Dietterich1998] Thomas G. Dietterich. 1998. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10:1895–1923.
- [Esuli and Sebastiani2006] Andrea Esuli and Fabrizio Sebastiani. 2006. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In *Proceedings of 5th Conference on Language Resources and Evaluation LREC 2006*, pages 417–422.
- [Haniewicz et al.2013] Konstanty Haniewicz, Wojciech Rutkowski, Magdalena Adamczyk, and Monika Kaczmarek. 2013. Towards the lexicon-based sentiment analysis of Polish texts: Polarity lexicon. In Costin Bădică, Ngoc Thanh Nguyen, and Marius Brezovan, editors, *Computational Collective Intelligence. Technologies and Applications: 5th International Conference, ICCCI 2013, Craiova, Romania, September 11-13, 2013, Proceedings*, pages 286–295. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Haniewicz et al.2014] Konstanty Haniewicz, Monika Kaczmarek, Magdalena Adamczyk, and Wojciech Rutkowski. 2014. Polarity lexicon for the Polish language: Design and extension with random walk algorithm. In J. Swiątek, A. Grzech, P. Swiątek, and J. M. Tomczak, editors, *Advances in Systems Science: Proc. of the Int. Conf. on Systems Science 2013 (ICSS 2013)*, pages 173–182. Springer.
- [Kulisiewicz et al.2015] Marcin Kulisiewicz, Tomasz Kajdanowicz, Przemysław Kazienko, and Maciej Piasecki, 2015. *On Sentiment Polarity Assignment in the Wordnet Using Loopy Belief Propagation*, pages 451–462. Springer International Publishing, Cham.
- [Liu2015] Bing Liu. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press, 1.
- [Mahyoub et al.2014] Fawaz H.H. Mahyoub, Muazam A. Siddiqui, and Mohamed Y. Dahab. 2014. Building an arabic sentiment lexicon using semi-supervised learning. *Journal of King Saud University - Computer and Information Sciences*, 26(4):417–424. Special Issue on Arabic NLP.
- [Maks and Vossen2011] Isa Maks and Piek Vossen. 2011. Different approaches to automatic polarity annotation at synset level. In *Proceedings of the First International Workshop on Lexical Resources*, pages 62–69.
- [Maks et al.2014] Isa Maks, Ruben Izquierdo, Francesca Frontini, Rodrigo Agerri, Piek Vossen, and Andoni Azpeitia. 2014. Generating polarity lexicons with wordnet propagation in 5 languages. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, may. European Language Resources Association (ELRA).
- [Maziarz et al.2016] Marek Maziarz, Maciej Piasecki, Ewa Rudnicka, Stan Szpakowicz, and Pawel Kedzia. 2016. plWordNet 3.0 – a comprehensive lexical-semantic resource. In *COLING 2016, 26th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, December 11-16, 2016, Osaka, Japan*, pages 2259–2268.
- [Miller1995] George A. Miller. 1995. WordNet: A lexical database for english. *Commun. ACM*, 38(11):39–41, November.
- [Misiaszek et al.2013] Andrzej Misiaszek, Tomasz Kajdanowicz, Przemysław Kazienko, and Maciej Piasecki, 2013. *Relational Propagation of Word Sentiment in WordNet*, pages 137–140. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Piek Vossen and VanderVliet2008] Roxane Segers Piek Vossen, Isa Maks and Hennie VanderVliet. 2008. Integrating lexical units, synsets and ontology in the cornetto database. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may. European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2008/>.
- [Stone1966] Philip J. Stone. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. The MIT Press.
- [Strapparava and Valitutti2004] Carlo Strapparava and Alessandro Valitutti. 2004. WordNet-Affect: An affective extension of WordNet. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 1083–1086.
- [Zaško-Zielińska et al.2015] Monika Zaško-Zielińska, Maciej Piasecki, and Stan Szpakowicz. 2015. A large wordnet-based sentiment lexicon for Polish. In Ruslan Mitkov, Galia Angelova, and Kalina Boncheva, editors, *Proceedings of the International Conference Recent Advances in Natural Language Processing – RANLP'2015*, pages 721–730, Hissar, Bulgaria. INCOMA Ltd. Shoumen, BULGARIA.