Propagation of emotions, arousal and polarity in WordNet using Heterogeneous Structured Synset Embeddings

Jan Kocoń[*], **Arkadiusz Janz**[*], **Monika Riegel**[†], **Małgorzata Wierzba**[†],
Artur Marchewka[†], Agnieszka Czoska[‡], Damian Grimling [‡],
Barbara Konat[‡¡], Konrad Juszczyk[‡¡], Katarzyna Klessa[¡], Maciej Piasecki[*]

[*]Wroclaw University of Science and Technology

[†] Laboratory of Brain Imaging, Nencki Institute of Experimental Biology of Polish Academy of Sciences

[‡]W3A.PL Sp. z o.o., ¡Adam Mickiewicz University

{jan.kocon,arkadiusz.janz,maciej.piasecki}@pwr.edu.pl,

{a.marchewka,m.riegel,m.wierzba}@nencki.gov.pl,

{agnieszka,damian,barbara,konrad}@sentimenti.pl,klessa@amu.edu.pl

Abstract

In this paper we present a novel method for emotive propagation in a wordnet based on a large emotive seed. We introduce a sense-level emotive lexicon annotated with polarity, arousal and emotions. The data were annotated as a part of a large study involving over 20,000 participants. A total of 30,000 lexical units in Polish WordNet were described with metadata, each unit received about 50 annotations concerning polarity, arousal and 8 basic emotions, marked on a multilevel scale. We present a preliminary approach to propagating emotive metadata to unlabeled lexical units based on the distribution of manual annotations using logistic regression and description of mixed synset embeddings based on our Heterogeneous Structured Synset Embeddings.

1 Introduction

Rapid growth of interest in sentiment analysis comes from its vast potential in automatic detection of subjectivity (whether the text expresses a subjective opinion rather than an objective fact) and polarity (whether the expressed opinion is positive, negative, or neutral) in large amount of textual data. For instance, sentiment analysis systems proved useful in automatic analysis of many different kinds of textual data, such as emails, tweets, blogs, reviews, newspaper headlines or novels (Dodds et al., 2015; Mohammad, 2016). Whereas introduction of advanced computational methods (e.g. machine learning) to natural language processing resulted in the development of sentiment analysis methodologies, the scarcity of high quality and large scale data sources greatly constrains their usage.

Numerous attempts were made to annotate words in terms of emotions for various languages (Riegel et al., 2015). However, such datasets are typically limited in size and consist of several thousands of words, while natural lexicons are known to be much bigger. Since annotations are provided manually by either qualified experts (usually 2-3 independent annotators) or a group of naive participants, data collection is typically very expensive in terms of time and money. Therefore, most of the available resources describe word meanings in terms of polarity, without further distinguishing various emotion categories attributed to them.

In emotion research, words are usually characterized according to two dominant theoretical accounts on the nature of emotion: dimensional account and categorical account. According to the first one proposed in (Russell and Mehrabian, 1977), each emotional state can be represented by its location in a multidimensional space, where valence or polarity (ranging from negativity to positivity) and arousal (from low to high) explain most of the observed variance. A competing account distinguishes several basic categories of emotional states, with more complex, subtle emotion states emerging as their combination. There have been various interpretations of the basic emotions concept, and different numbers of emotion categories were proposed by different theories, with (Ekman, 1992) and (Plutchik, 1982) gaining most recognition in the scientific community.

In this work, we used a large dataset described in (Kocoń et al., 2019a), containing metadata for a total of over 30000 word meanings from Polish WordNet (Piasecki et al., 2009), annotated

in terms of polarity, arousal, as well as 8 basic emotion categories (i.e. joy, sadness, trust, disgust, fear, anger, surprise, anticipation). Here, we present a novel propagation approach to automatically extend the original dataset by deriving emotion metadata for lexical units and synsets that are not present in this dataset. If effective, our approach could alleviate the problem of data scarcity and facilitate the widespread use of sentiment analysis in various applications including but not restricted to artificial intelligence, computational linguistics, psychology or business.

2 Dataset description

In the Sentimenti database (Kocoń et al., 2019a), a total of over 20,000 unique respondents (with approximately equal number of male and female participants) was sampled from Polish population. Multiple demographical characteristics such as: sex, age, native language, place of residence, education level, marital status, employment status, political beliefs and income were controlled. The annotation schema was based on the procedures most widely used in previous studies, aiming to create the first standardized datasets of Polish words characterized in terms of emotion (NAWL, (Riegel et al., 2015); NAWL BE, (Wierzba et al., 2015); plWordNet-emo (Zaśko-Zielińska et al., 2015; Janz et al., 2017)). Thus, we collected annotations of valence (polarity), arousal, as well as eight emotion categories: joy, sadness, trust, disgust, fear, anger, surprise and anticipation. By combining simple annotation schema with crowd annotation, we were able to effectively acquire large amount of data and preserve its high quality at the same time.

The total number of over 30,000 word meanings was annotated, with each meaning ranked at least 50 times on each scale. Moreover, in a follow-up study a total number of over 7,000 texts (short phrases or paragraphs of text) were annotated in the same way, with each text assessed at least 25 times on each scale. Before attempting the assessment task, subjects were instructed to rank word meanings rather than words, as well as encouraged to indicate their immediate, spontaneous reactions. Participants had unlimited time to complete the task and they were able to quit the assessment session at any time and resume their work later on. The source of texts were reviews from

two domains: *medicine*[1] (2000 reviews) and *hotels*[2] (2000 reviews). Due to difficulties in observing neutral reviews in the selected sources, we have chosen them from websites describing medical information[3] (500 paragraphs) and the hotel industry[4] (500 paragraphs). We also selected phrases using *lexico-semantic-syntactic patterns* (LSS) manually created by linguists to capture one of the four effects affecting sentiment: *increase, decrease, transition, drift*. Most of these phrases belong to the previously mentioned subject areas. The source for the remaining phrases were Polish WordNet glosses (Piasecki et al., 2009).

2.1 Data conversion

We decided to treat the problem of emotive propagation as a multilabel classification task, where the individual lexical units are classified considering all emotive categories. Eight basic emotions and arousal were annotated on a scale of integers from range $[0, 4]$. The valence was annotated using $[-3, 3]$ scale. To perform the classification task a proper conversion schema should be applied. For most of emotive dimensions we used a simple averaging strategy, where the final score is an average value of all assigned scores, normalized to the range $[0, 1]$. In the case of valence scores we divided the annotations into two separate groups: positive scores (Valence$_p$) and negative scores (Valence$_n$). This division results from the fact that some texts have mixed annotations, both positive and negative. To keep the original distribution of valence scores we decided to use a separate approach (see Algorithm 1). The positive scores were separated from negative ones to measure the degree of positive valence (valence$_p$) and negative valence (valence$_n$). With this approach a single lexical unit obtains two normalized valence scores.

We decided to partition all scores for each dimension into two clusters using k-means clustering (Hartigan and Wong, 1979). We assign a label representing a membership of an individual lexical unit to specific emotive category if the final score of a lexical unit in the dimension representing that category is greater than the threshold determined by k-means. Each lexical unit can be described by multiple categories, thus we might obtain multiple

---

[1] www.znanylekarz.pl
[2] pl.tripadvisor.com
[3] naukawpolsce.pap.pl/zdrowie
[4] hotelarstwo.net, www.e-hotelarstwo.com

Algorithm 1 Estimating the average value of positive and negative valence for a single lexical unit.

Require: V: list of all valence scores; $m = 3$: maximum absolute value of polarity
Ensure: Pair (p, n) where p is average positive valence, and n is average negative valence
  1: $(p, n) = (0, 0)$
  2: **for** $v \in V$ **do**
  3:     **if** $v < 0$ **then** $n = n + |v|$ **else** $p = p + v$
      return $(p \div (|V| \cdot m), n \div (|V| \cdot m))$

---

labels assigned to a single lexical unit.

2.2 Polarity transfer from units to synsets

On the basis of the previous work where we analysed the contemporary annotation of plWordNet to see how diverse synsets are in terms of lexical units valence, we assumed that we can average all dimensions of lexical units (which belong to synset A) separately to obtain the metadata description of synset A. Previously acquired statistics show that synsets are strongly homogeneous in terms of the units valence, so we decided to move annotations from unit-level to synset-level in that way (Kocoń et al., 2018a; Kocoń et al., 2018b).

3 Emotive Propagation

In this study, we decided to follow the idea presented in (Kocoń et al., 2018a; Kocoń et al., 2018b). The idea is to apply a semi-supervised learning on a large seed of emotively annotated synsets. The seed is used to train a classifier and then to predict the polarity categories for unlabeled synsets in a close vicinity of labeled ones. Starting from the labeled synsets we visit their neighbors and annotate them iteratively by applying our classifier. The propagation process ends when all synsets become annotated.

In (Kocoń et al., 2018b) the authors proposed a rich set of wordnet-based features to describe the synsets representing an initial seed for emotive propagation. For every synset existing in a seed they extracted the features capturing the structure of its neighborhood by taking into account the neighboring synsets (with their relative location) and assigned polarities. The features were generated on a basis of a template being defined in terms of 4 feature variables. A single feature is generated by initializing the variables in the template with a specific combination of possible values. The variables in the template are defined as follows:

- *Relation* – one of the 13 most common WordNet relations,

- *Direction* – the direction of the relation,

- *Element* – a type of an element used to construct a *bag-of-words* model; two types of *elements* were used: *synset_ID* (any number) and *synset_polarity* (one of the following numbers: $-1, 0, 1$; it represents 3 polarity classes: *negative, neutral, positive*),

- *Level* – a distance (number of hops) between the initial synset being described and its neighbors, e.g. the synsets of *second level* means neighboring synsets in a distance of two hops (excluding the synsets of *first level*).

Extracted features were converted to a *bag-of-words* model, where the elements of a bag were representing the synsets or their polarities. Then the authors used these features as a signal for a classifier to decide whether a given synset should be positive, negative, neutral or ambiguous.

Such an approach generated vectors of very large dimensions, thus it increases the overall propagation time. The classification procedure was time-consuming because the process of feature generation produced high-dimensional data.

3.1 Heterogeneous Structured Synset Embeddings

To reduce the dimensionality of the input feature space, we decided to build upon the methods designed for embedding the lexical knowledge bases. The main aim is to produce a meaningful vector space representation of concepts existing in a knowledge base by capturing their lexico-semantic properties and embedding the structure of their neighborhood. In our case the concepts are represented by synsets and we utilize the structure of a wordnet to construct synset embeddings. Our approach is based on the skip-gram model (Mikolov et al., 2013) which takes as its input a large textual corpus and produces a distributional representation of words by capturing the neighboring words appearing in a small context window. The main assumption is that the words sharing the similar context should have similar vector space representations. The neural network based on skip-gram architecture learns the

vector space representations of words (existing in a corpus) by minimizing the loss in the task of *predicting the context words given the input word.*

To produce synset embeddings we adapted the solution presented in (Goikoetxea et al., 2015). The authors generated an artificial wordnet-based corpus to train a skip-gram model by performing multiple random walks on a wordnet. The input corpus consisted of synset identifiers generated during random walking process. Previous solutions were limited only to synset links and did not include the information about relation types. We decided to expand this idea by including the lexical units and their links while generating the corpus with a random walk. Thus, we might obtain a corpus with artificial words (elements) represented by the identifiers of synsets, lexical units, and the identifiers of relation types. In the case of relation types we decided to differentiate the identifiers depending on the type of linked concepts, e.g. the identifiers starting with the rSS prefix represent the relations between synsets while rSL (or rLS) represent the links between synsets and lexical units. The elements with the rSS prefix have an additional identifier representing a specific type of wordnet relation (e.g. 10 represents *hyponymy*). The additional information about the types of links should lead to obtaining a more heterogeneous and accurate embeddings of synsets. A short random walk of length 13 can be represented as the following sequence of elements:

```
s_7078349 -> rSS_136 -> s_60485 -> rSL
-> l_85957 -> rLS -> s_60485 -> rSS_10
-> s_22456 -> rSS_11 -> s_55576 -> rSS_11
-> s_55575 -> rSS_11 -> s_7077974
-> rSS_10 -> s_55575 -> rSL -> l_79892
-> rLL_3425 -> l_10483 -> rLS -> s_3974
-> rSS_11 -> s_7077977
```

To generate the embeddings, we used a popular FastText method (Bojanowski et al., 2017; Joulin et al., 2017). This method was used in many different NLP tasks, especially with applications to sentiment analysis e.g. hate speech detection (Badjatiya et al., 2017), sentiment polarity recognition, emotion and sarcasm identification (Felbo et al., 2017), aspect-based sentiment analysis in social media (Wojatzki et al., 2017), text classification on multiple sentiment datasets (Joulin et al., 2017).

3.2 Emotive classification

The knowledge base embeddings alone might be insufficient to successfully solve downstream tasks due to the lack of contextual information

connecting these embeddings with real world data. To prepare a more meaningful and contextual representation for our emotive classifier we decided to augment our model with plain word embeddings. In (Kocoń et al., 2018a) the authors showed that the size and the quality of training corpora might affect the overall performance in downstream tasks. They also tested several parameter settings of word embeddings for Polish language using the implementation of CBOW and skip-gram methods provided with FastText (Bojanowski et al., 2017). These models are available under an open license in the CLARIN-PL project DSpace repository[5]. With these embeddings, the best results were obtained in two NLP tasks: recognition of temporal expressions (Kocoń and Gawor, 2018; Kocoń et al., 2019b) and recognition of named entities (Marcińczuk et al., 2018).

We used the same model to produce a complementary feature space for the task of polarity prediction. To prepare a complementary embedding space for HSSE we averaged the embeddings of lemmas linked with the synsets through their lexical units. For each synset in a seed we computed the averaged embedding of its lemmas and concatenated it with HSSE embedding.

Our emotive classifier is represented as an ensemble of binary classifiers, each one predicting one of 11 emotive categories. Each classifier in this ensemble was trained on a seed of synsets representing a specific emotive category with its positive and negative examples.

4 Results and Discussion

Comparing the F1 scores of the model in a task of valence, arousal and emotion propagation, we can observe that the valence propagation performs better than the propagation of most emotive dimensions (Table 1). The best results are obtained for *trust* (F1-score: 77.48%) and anticipation (F1-score: 74.94%). F-score for all dimensions except *disgust* are above 63%.

5 Conclusions

The results of the propagation task suggest that the novel method presented here can be useful for both valence and emotions datasets. Heterogeneous Structured Synset Embeddings allow for effective scaling up of annotated datasets.

---

[5]https://clarin-pl.eu/dspace/handle/11321/606

Dim.	P[%]	R[%]	F[%]
$Valence_p$	70.18	81.17	75.27
$Valence_n$	65.57	85.26	74.12
Arousal	66.74	75.76	70.97
Joy	59.27	80.56	68.29
Surprise	59.55	69.37	64.07
Anticip.	70.58	79.12	74.94
Trust	74.86	80.46	77.48
Sadness	59.18	83.16	69.13
Anger	56.99	84.14	67.93
Fear	51.98	81.40	63.43
Disgust	45.50	82.65	58.71

Table 1: Precision, recall and F1-score for Valence, Arousal and Emotion propagation.

The paper also describes a large dataset: "Sentimenti" which covers more than 30,000 word-meanings in Polish annotated with 8 basic emotions as well as polarity and arousal.

The data gathered in the psycholinguistic study were used to enhance affective annotation of Polish WordNet (Piasecki et al., 2009). Tthe emotive metadata were propagated to unlabeled lexical units, enabling emotive categorization of the whole WordNet. Categorization results proved that the ML methods trained on the data enriched with detailed description of synset features and relations between lexical units and synsets resulted in effective models for emotional metadata propagation.

Such metadata propagation methods are successful only when based on large data sets and multilevel annotations of a given WordNet. In the current study, we showed that their effectiveness can be high also for very subjective emotional features, especially valence propagation. Emotional metadata propagation for Polish proved to have a high accuracy for most of the emotion values. However, a question remains why some of the emotion values were attributed less effectively - whether it was the lack of input data or it is a property of those emotions expressed verbally to be less evident and more dispersed.

Acknowledgments

References

[Badjatiya et al.2017] Pinkesh Badjatiya, Shashank Gupta, Manish Gupta, and Vasudeva Varma. 2017. Deep learning for hate speech detection in tweets. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 759–760. International World Wide Web Conferences Steering Committee.

[Bojanowski et al.2017] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

[Dodds et al.2015] Peter Sheridan Dodds, Eric M Clark, Suma Desu, Morgan R Frank, Andrew J Reagan, Jake Ryland Williams, Lewis Mitchell, Kameron Decker Harris, Isabel M Kloumann, James P Bagrow, Karine Megerdoomian, Matthew T McMahon, Brian F Tivnan, and Christopher M Danforth. 2015. Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences of the United States of America*, 112(8):2389–94.

[Ekman1992] Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion*, 6(3-4):169–200.

[Felbo et al.2017] Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1615–1625.

[Goikoetxea et al.2015] Josu Goikoetxea, Aitor Soroa, and Eneko Agirre. 2015. Random walks and neural network language models on knowledge bases. In *Proceedings of the 2015 conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies*, pages 1434–1439.

[Hartigan and Wong1979] John A Hartigan and Manchek A Wong. 1979. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108.

[Janz et al.2017] Arkadiusz Janz, Jan Kocoń, Maciej Piasecki, and Zaśko-Zielińska Monika. 2017. plWordNet as a Basis for Large Emotive Lexicons of Polish. In *LTC'17 8th Language and Technology Conference*, Poznań, Poland, November. Fundacja Uniwersytetu im. Adama Mickiewicza w Poznaniu.

[Joulin et al.2017] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2017. Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 427–431, Va-

lencia, Spain, April. Association for Computational Linguistics.

[Kocoń and Gawor2018] Jan Kocoń and Michał Gawor. 2018. Evaluating KGR10 Polish Word Embeddings in the Recognition of Temporal Expressions Using BiLSTM-CRF. *Schedae Informaticae*, 2018(Volume 27).

[Kocoń et al.2018a] Jan Kocoń, Arkadiusz Janz, and Maciej Piasecki. 2018a. Classifier-based Polarity Propagation in a Wordnet. In *Proceedings of the 11ᵗʰ International Conference on Language Resources and Evaluation (LREC'18)*.

[Kocoń et al.2018b] Jan Kocoń, Arkadiusz Janz, and Maciej Piasecki. 2018b. Context-sensitive Sentiment Propagation in WordNet. In *Proceedings of the 9ᵗʰ International Global Wordnet Conference (GWC'18)*.

[Kocoń et al.2019a] Jan Kocoń, Arkadiusz Janz, Piotr Miłkowski, Monika Riegel, Małgorzata Wierzba, Artur Marchewka, Agnieszka Czoska, Damian Grimling, Barbara Konat, Konrad Juszczyk, Katarzyna Klessa, and Maciej Piasecki. 2019a. Recognition of emotions, valence and arousal in large-scale multi-domain text reviews. In Zygmunt Vetulani and Patrick Paroubek, editors, *Human Language Technologies as a Challenge for Computer Science and Linguistics*. Wydawnictwo Nauka i Innowacje, Poznań, Poland, May.

[Kocoń et al.2019b] Jan Kocoń, Marcin Oleksy, Tomasz Bernaś, and Michał Marcińczuk. 2019b. Results of the PolEval 2019 Shared Task 1: Recognition and Normalization of Temporal Expressions. *Proceedings of the PolEval2019 Workshop*.

[Marcińczuk et al.2018] Michał Marcińczuk, Jan Kocoń, and Michał Gawor. 2018. Recognition of Named Entities for Polish-Comparison of Deep Learning and Conditional Random Fields Approaches. In Maciej Ogrodniczuk and Łukasz Kobyliński, editors, *Proceedings of the PolEval 2018 Workshop*, pages 77–92. Institute of Computer Science, Polish Academy of Science.

[Mikolov et al.2013] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

[Mohammad2016] Saif M. Mohammad. 2016. Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In Herb Meiselman, editor, *Emotion Measurement*. Elsevier.

[Piasecki et al.2009] Maciej Piasecki, Bernd Broda, and Stanislaw Szpakowicz. 2009. *A wordnet from the ground up*. Oficyna Wydawnicza Politechniki Wrocławskiej Wrocław.

[Plutchik1982] Robert Plutchik. 1982. A psychoevolutionary theory of emotions. *Social Science Information*, 21(4-5):529–553.

[Riegel et al.2015] Monika Riegel, Małgorzata Wierzba, Marek Wypych, Łukasz Żurawski, Katarzyna Jednoróg, Anna Grabowska, and Artur Marchewka. 2015. Nencki Affective Word List (NAWL): the cultural adaptation of the Berlin Affective Word List–Reloaded (BAWL-R) for Polish. *Behavior Research Methods*, 47(4):1222–1236.

[Russell and Mehrabian1977] James A Russell and Albert Mehrabian. 1977. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3):273 – 294.

[Wierzba et al.2015] Małgorzata Wierzba, Monika Riegel, Marek Wypych, Katarzyna Jednoróg, Paweł Turnau, Anna Grabowska, and Artur Marchewka. 2015. Basic emotions in the Nencki Affective Word List (NAWL BE): New method of classifying emotional stimuli. *PLOS ONE*, 10(7):1–16.

[Wojatzki et al.2017] Michael Wojatzki, Eugen Ruppert, Sarah Holschneider, Torsten Zesch, and Chris Biemann. 2017. Germeval 2017: Shared task on aspect-based sentiment in social media customer feedback. *Proceedings of the GermEval*, pages 1–12.

[Zaśko-Zielińska et al.2015] Monika Zaśko-Zielińska, Maciej Piasecki, and Stan Szpakowicz. 2015. A large wordnet-based sentiment lexicon for Polish. In *Proceedings of the International Conference Recent Advances in Natural Language Processing*, pages 721–730.