

The Role of the WordNet Relations in the Knowledge-based Word Sense Disambiguation Task

Kiril Simov

Alexander Popov

Petya Osenova

Linguistic Modeling Department

IICT, Bulgarian Academy of Sciences

{kivs|alex.popov|petya}@bultreebank.org

Abstract

In this paper we present an analysis of different semantic relations extracted from WordNet, Extended WordNet and SemCor, with respect to their role in the task of knowledge-based word sense disambiguation. The experiments use the same algorithm and the same test sets, but different variants of the knowledge graph. The results show that different sets of relations have different impact on the results: positive or negative. The beneficial ones are discussed with respect to the combination of relations and with respect to the test set. The inclusion of inference has only a modest impact on accuracy, while the addition of syntactic relations produces stable improvement over the baselines.

1 Introduction

Knowledge-based methods for Word Sense Disambiguation (WSD) are attractive to the NLP community because they do not require manually annotated corpora. On the other hand, these methods are not considered completely unsupervised, because they do need information about senses of words in texts, and about the relations that hold between them, represented in the form of a directed or undirected graph, called **knowledge graph** (KG). The most frequently used knowledge graph is based on WordNet (WN) (Fellbaum, 1998) or Extended WordNet (XWN) (Mihalcea and Moldovan, 2001), where synsets constitute the vertices of the graph and relations between synsets are represented as edges within it. Simov et al. (2015) provided evidence that the addition of linguistically motivated semantic relations to the KG improves the performance of Knowledge-based WSD (KWSD). In the current work we perform an analysis of the various semantic relations

in WN and XWN knowledge graphs. The analysis is performed via experiments with different subgraphs that include only some of the semantic relations in WN and XWN. Some of the relation types allow for inference to be applied over them. Thus, inferred semantic relations have been included in some of KGs as well. The experiments were performed on the manually annotated SemCor corpus (Miller et al., 1993). In order to test the semantic relations extracted from the syntactically annotated corpus, the same was divided into four parts. We used three of the divisions for the extraction of new relations and one part for testing.

The structure of the papers is as follows: the next section discusses related work on the topic. Section 3 describes the experimental setup. Section 4 focuses on the experiments with the semantic relations in WordNet. Section 5 presents the experiments with the semantic relations in Extended WordNet. Section 6 gives an overview of the experiments with syntactic relations. Section 7 concludes the paper.

2 Related Work

Knowledge-based systems for WSD have proven to be a good alternative to supervised systems, which require large amounts of manually annotated training data. In contrast, knowledge-based systems require only a knowledge base and no additional corpus-dependent information. An especially popular knowledge-based disambiguation approach has been the use of successful graph-based algorithms known under the name of "Random Walk on Graph" (Agirre et al., 2014). Most methods exploit variants of the PageRank algorithm (Brin and Page, 2012). Agirre and Soroa (2009) apply a variant of the algorithm to Word Sense Disambiguation by translating WordNet into a knowledge graph in which the synsets are represented as vertices and the relations between them are represented as edges between the ver-

tices. Calculating the PageRank vector \mathbf{Pr} is accomplished through solving the equation:

$$\mathbf{Pr} = cM\mathbf{Pr} + (1 - c)\mathbf{v} \quad (1)$$

where M is an $N \times N$ transition probability matrix (N being the number of vertices in the graph), c is the damping factor and \mathbf{v} is an $N \times 1$ vector. In the traditional, static version of PageRank the values of \mathbf{v} are all equal ($1/N$), which means that in the case of a random jump each vertex is equally likely to be selected. Modifying the values of \mathbf{v} effectively changes these probabilities and thus makes certain vertices more important. The version of PageRank for which the values in \mathbf{v} are not uniform is called *Personalized PageRank*.

The words in the text that are to be disambiguated are inserted as nodes in the knowledge graph and are connected to their potential senses via directed edges (by default, a context window of at least 20 words is used). These newly introduced nodes serve to inject initial probability mass (via the \mathbf{v} vector) and thus to make their associated sense nodes especially relevant in the knowledge graph. Applying the Personalized PageRank algorithm iteratively over the graph determines the most appropriate sense for each ambiguous word. The method has been boosted by the addition of new relations and by developing variations and optimizations of the algorithm. It has also been applied to the task of NED (Agirre et al., 2015).

The success of KWSD approaches apparently depends on the quality of the knowledge graph – whether the knowledge represented in terms of nodes and relations (edges/links) between them is sufficient for the algorithm to pick the correct senses of ambiguous words. Several extensions of the knowledge graph, constructed on the basis of WordNet, have been proposed and implemented. In (Simov et al., 2015), semantic and syntactic relations from the sense annotated BulTreeBank have been extracted and the algorithm has been applied to Bulgarian data. In order to do that, the treebank was first annotated with synsets from the BulTreeBank WordNet¹, aligned to the Princeton WordNet. The word forms annotated with senses at this point are 69,333, consisting of nouns and verbs. Out of these, 12,792 sense-tagged word

¹The Core WordNet is freely available at: <http://compling.hss.ntu.edu.sg/omw/>. The extended one will be released soon. For more details about the sense annotated BulTreeBank, see (Popov et al., 2014).

forms have been used for testing, and the rest have been used for relation extraction.

The WordNet ontological relations that have been used are 252,392, and the relations derived from the synset glosses are 419,387. Additionally, the following relations have been extracted: inferred hypernymy relations; syntactic relations from the gold corpus; extended syntactic relations; domain relations from WordNet. Thus, 590,272 new relations have been added. The newly added relations introduce syntagmatic information into the graph, which was originally constructed out of paradigmatic relations. The results from the experiments with paradigmatic relations alone (done on the the whole corpus) show highest accuracy (0.551) for the combination of: WordNet relations + relations from the glosses + inferred hypernymy relations + domain relations of the kind synset-to-synset + domain hierarchy relations. The results from the experiments with mixed – paradigmatic and syntagmatic – relations (done on a test portion of one fourth of the corpus) show highest accuracy (0.656) for the combination of: WordNet relations + relations from XWN + inferred hypernymy relations + dependency relations from the golden corpus + extended dependency relations starting from one level up + domain relations of the kind synset-to-synset + domain hierarchy relations.

Kdzia et al. (2014) present work on WSD for Polish using the Polish WordNet, extended with relations between semantically similar words. The authors use the Measure of Semantic Relatedness which assigns a numerical value to pairs of words. This numerical value reflects the degree of closeness between two words. For each word w_i , a list of most closely related words w_j is constructed (length of the list is 20). Then the synsets that contain w_i and synsets containing some of w_j are connected with new links. The evaluation, based on the extended knowledge graph, shows improvement on the sentence level.

3 Experimental Set-Up

The experiments presented here were carried out with the UKB² tool, which provides graph-based methods for Word Sense Disambiguation and measuring lexical similarity. The tool uses the Personalized PageRank algorithm, described in Agirre and Soroa (2009). It builds a knowledge graph over a set of relations that can be induced

²<http://ixa2.si.ehu.es/ukb/>

from different types of resources, such as WordNet or DBPedia; then it selects a context window of open class words and runs the algorithm over the graph. There is an additional module called NAF UKB³ that can be used to run UKB with input in the NAF format⁴ and to obtain output structured in the same way, only with added word sense information. For compatibility reasons, NAF UKB was used to perform the experiments reported here; the input NAF document contains in its "term" nodes lemma and POS information, which is necessary for the running of UKB. We have used the UKB default settings, i.e. a context window of 20 words that are to be disambiguated together, 30 iterations of the Personalized PageRank algorithm.

The UKB tool requires two resource files to process the input file. One of the resources is a dictionary file with all lemmas that can be possibly linked to a sense identifier. In our case, WordNet-derived relations were used as our knowledge base; consequently, the sense identifiers are WordNet IDs. For instance, a dictionary line compiled from WordNet synsets looks like this:

```
predicate 06316813-n:0 06316626-n:0
01017222-v:0 01017001-v:0 00931232-v:0
```

It comprises of a lemma followed by the sense identifiers it can be associated with. Each ID consists of eight digits followed by a hyphen and a label referring to the POS category of the word. Finally, a number following a colon indicates the frequency of the word sense, calculated on the basis of a tagged corpus. When a lemma from the dictionary has occurred in the analysis of the input text, the tool assigns all associated word senses to the word form in the context and attempts to disambiguate its meaning among them. The Bulgarian dictionary comprises of all the lemmas of words annotated with WordNet senses in the BTB. It has 8,491 lemmas mapped to 6,965 unique word senses.

The second resource file required for running the tool is the set of relations that is used to construct the knowledge graph over which Personalized PageRank is run. The distribution of the tool provides data (dictionary and relation files) for WordNet 1.7 and 3.0. Since the BTB has been annotated with word senses from WordNet 3.0, the resource files for version 3.0 have been used in our experiments. The distribution of UKB comes with

³https://github.com/asoroa/naf_ukb

⁴<http://www.newsreader-project.eu/files/2013/01/techreport.pdf>

a file containing the standard lexical relations defined in WordNet, such as hypernymy, meronymy, etc., as well as with a file containing relations derived on the basis of common words found in the synset glosses, which have been manually disambiguated. The format of the relations in the knowledge graph is as follows:

```
u:SynSetId01 v:SynSetId02 s:Source d:w
```

where `SynSetId01` is the identifier of the first synset in the relation, `SynSetId02` is the identifier of the second synset, `Source` is the source of the relation, and `w` is the weight of the relation in the graph. In the experiments reported in the paper, the weight of all relations is set to 0. Here is one concrete example:

```
u:01916925-n v:02673969-a s:30g1c d:0
```

All the experiments use the same algorithm and the same test data. Only the knowledge graph differs in the different cases, as it is generated out of various sets of relations.

The experiments, reported in Table 1, are considered baselines for the two semantically annotated corpora: the first 49 documents of SemCor (about 1/4 of the data) and the three selected documents from BulTreeBank (about 1/4 of the data). The baseline results include WordNet relations (WN), gloss-derived relations (GL) and the combination of WN and GL — WNG:

<i>KG</i>	<i>SemCor</i>	<i>BTB</i>
WN	49.24	51.72
GL	51.48	47.02
WNG	58.83	53.82

Table 1: Experimental results when using the original knowledge graphs (WN, GL, WNG) on the two test corpora.

Some considerations are in order. It is apparent that the results for the English corpus increase monotonically, while for the Bulgarian one they are non-monotonic. Also, the combined WordNet and gloss-derived relations increase the SemCor results a lot more than the BTB ones. This probably reflects the fact that these are, after all, glosses in English and they capture better meanings encoded in the English corpus.

4 Experiments with Semantic Relations in WordNet

The WordNet-based KG (WN) has been constructed out of the relations in the Princeton WordNet (PWN3.0). PWN3.0 groups together words in synsets, which we consider as concepts, and thus as units. The relation types possible between the different synsets are 16. In our experiments we separated the relations in WN into 16 sets of relations corresponding to the relations in PWN3.0:

1. **WN-Hyp** (*hypernymy*) **89089**. (N-N), (V-V)⁵.
2. **WN-Ant** (*antonymy*) **8689**. (A-A), (N-N), (R-R), (V-V).
3. **WN-At** (*attribute relation between noun and adjective*) **886**. (N-A), (A-N).
4. **WN-Cls** (*a member of a class*) **9420**. (A-N), (N-N), (R-N), (V-N).
5. **WN-Cs** (*cause*) **192**. (V-V).
6. **WN-Der** (*derivational morphology*) **74644**. (A-N), (N-A), (N-N), (N-V).
7. **WN-Ent** (*entailment*) **408**. (V-V).
8. **WN-Ins** (*instance*) **8576**. (N-N).
9. **WN-Mm** (*member meronymy*) **12293**. (N-N).
10. **WN-Mp** (*part meronymy*) **9097**. (N-N).
11. **WN-Ms** (*substance meronymy*) **797**. (N-N).
12. **WN-Per** (*pertains/derived from*) **8505**. (A-N), (R-A).
13. **WN-Ppl** (*participle of the verb*) **79**. (A-V).
14. **WN-Sa** (*additional information about the first word*) **3269**. (A-A), (V-V).
15. **WN-Sim** (*similar in meaning*) **21386**. (A-A).
16. **WN-Vgp** (*similar in meaning verb synsets*) **1725**. (V-V).

These classes differ in the type of semantic relations they represent, the number of relations in each class, the parts-of-speech of the words in the synsets that are connected by the relation. Obviously, isolated vertices do not play a role in the disambiguation process. Thus, if we exploit only relations between nouns, we cannot expect that the system could select appropriate senses

⁵Here we present the combination of synsets in each relation as parts-of-speech. The parts-of-speech are: A — adjective, N — noun, R — adverb, and V — verb. Also we present the number of links for the relation in WordNet.

for other parts-of-speech. Nevertheless, we performed some experiments with only some of the relations in order to have a basis for comparison with larger combinations. As a basic relation we consider the superordinate-subordinate relation (hypernymy), because it provides relations between the biggest groups of synsets: nouns and verbs. Thus, we assume that this set of relations always has to be used in the knowledge graph.

<i>KG</i>	<i>SemCor</i>	<i>BTB</i>
WN	49.24	51.72
GL	51.48	47.02
WNG	58.83	53.82
WN-Hyp	33.38	44.89
WN-Hyp+WN-Ant	39.79	47.55
WN-Hyp+WN-At	35.77	46.18
WN-Hyp+WN-Cls	34.12	46.11
WN-Hyp+WN-Cs	33.30	40.94
WN-Hyp+WN-Der	38.93	49.26
WN-Hyp+WN-Ent	33.09	44.29
WN-Hyp+WN-Ins	33.89	45.00
WN-Hyp+WN-Mm	33.42	44.61
WN-Hyp+WN-Mp	35.60	45.03
WN-Hyp+WN-Ms	33.32	45.00
WN-Hyp+WN-Per	39.62	47.29
WN-Hyp+WN-Ppl	33.29	40.57
WN-Hyp+WN-Sa	38.07	44.48
WN-Hyp+WN-Sim	42.71	44.49
WN-Hyp+WN-Vgp	33.96	41.11

Table 2: Experimental results when using the sets of relations from the WordNet knowledge graph on the two test corpora.

In Table 2 we present the results for combinations between the hypernymy relation and all other relations. The biggest improvement is observed for the combination **WN-Hyp+WN-Sim**. It shows 9 % of improvement over the **WN-Hyp** relation alone. In our view, the great difference is due to the different coverage of the relations over the synsets in WordNet. Hypernymy relation covers only noun and verb synsets, but not adjective and adverb synsets. Thus, a KG based only on hypernymy relation does not provide any knowledge about adjectives and adverbs. Additionally, it does not contain any knowledge about the interactions between verbs and nouns. The relations that improve over hypernymy ones in fact introduce knowledge about adjectives or interaction

across parts-of-speech. We have performed some more experiments in order to check whether we could exclude some relations without considerable loss. For instance, the combination of the following eight sets: **WN-Hyp + WN-Ant + WN-Der + WN-Per + WN-Sa + WN-Sim + WN-Mp + WN-Cls**, gives accuracy of **49.10 %** on the *SemCor* test corpus, which is **0.14 %** less than the accuracy obtained with the whole KG of WordNet. The results also show the differences between the corpora. BTB seems more compact with respect to sub-domains, while SemCor introduces a big variety of sub-domains. Also, it is mainly annotated with noun and verb synsets. Thus, the impact of the relations is different from the impact they have over the SemCor corpus.

The general conclusion from these experiments is that the addition of relations to the knowledge graph does not contribute monotonically to the accuracy of the KWSD. It shows that some of the relations in the original graph lower the accuracy.

In the next sections we report only experiments performed over SemCor corpus for brevity.

4.1 Inference over WordNet Relations

Under inference in our experiments we consider the application of rules, given relations in the knowledge graph, which produce new relations to be added to the knowledge graph. In this section we consider some rules applicable to the relations from WordNet. Having in mind that WordNet is not a fully formalized lexical database, we cannot expect that the inferences proposed below are always correct. The main inference rule is the hypernymy hierarchy inheritance: if some relation includes a noun as an argument, then the hyponyms of the noun also could be arguments in the relation. The situation is similar for verbs. Sometimes the appropriate inference includes their hypernyms.

1. **WN-Hyp.** The hypernymy relation is transitive. Thus, we could construct its transitive closure: if **doctor** is a hypernym of **surgeon** and **professional** is a hypernym of **doctor** then **professional** is a hypernym of **surgeon**. Similarly, for the verb hierarchy.
2. **WN-Ant.** Antonymy relations between adjectives and adverbs cannot participate in the inference, because there is no support in WordNet. For nouns and verbs it is possible, if we assume that the antonymy relation

means that corresponding synsets are disjoint. The disjointedness is preserved by the hyponymy relation: if we have two disjoint concepts, then their subconcepts are also disjoint. For example, **man** and **woman** do not have common instances. Then we could infer that **man** and **girl** are disjoint.

3. **WN-At.** The attributes of a noun usually can be inherited by its hyponyms. For example, **measure** as a quantity of something has attributes — **standard** and **nonstandard**. These attributes can be inherited by all kinds of measures like **time interval** and others.
4. **WN-Cls.** The general understanding of the relation *a member of a class* is that each hyponym of the **member** could be a member of each of the hypernyms of the **class**. For instance, **desktop publishing** is a member of **computer science** as a branch, but also it is a branch of **engineering**, which is a hypernym of computer science.
5. **WN-Cs.** The *cause* relation between verbs naturally allows for inference on both arguments — each hyponym of the first argument could be a cause for each hypernym of the second argument. The sets resulting from the inference on the first and second arguments are denoted with **WN-Cs1stVerbInfer** and **WN-Cs2ndVerbInfer**.
6. **WN-Der.** The derivational relation is quite diverse, connecting adjectives and nouns, nouns and nouns, and nouns and verbs. We consider this relation as denoting an event or a state in which the noun determines a participant of the event or a state. Thus, a noun can be substituted with its hyponyms, and a verb can be substituted with its hypernyms.
7. **WN-Ent.** If a verb entails another verb, then we assume that each hyponym of the first verb entails each hypernym of the second verb. The sets resulting from the inference on the first and second arguments are denoted with **WN-Ent1stVerbInfer** and **WN-Ent2ndVerbInfer**.
8. **WN-Ins.** An instance of a class is an instance of its super classes. Thus, we perform substitution of the second noun with its hypernyms.
9. **WN-Mm.** Each hyponym of a member of a set is a member of each hypernym of the set.

10. **WN-Mp.** The transitive closure over the part meronym relation is a feasible inference rule. In these experiments we do not perform it.
11. **WN-Ms.** Substitution with hyponyms of the substance noun is a feasible inference rule. Similarly to the previous relation, in these experiments we do not perform it.
12. **WN-Per.** Similarly to the derivational relation, we perform substitution with hyponyms on the noun synset.
13. **WN-Ppl.** We do not perform any inference for this relation.
14. **WN-Sa.** The additional information about the first word can be inherited by its hyponyms.
15. **WN-Sim.** We do not perform any inference for this relation.
16. **WN-Vgp.** Because the definition “verb synsets that are similar in meaning” allows for very wide interpretation, we do not perform any inference on this relation.

Some of the above inferences produce a huge amount of new relations, which prevents us from effectively experimenting with them. We have used the inference rules only partially. These experiments have been performed only on the SemCor test corpus. We consider only combinations in which the knowledge graphs of the original WordNet and the Extended WordNet are included as a basis. Table 3 presents some of the results. There are few cases in which the inferred new relations add accuracy above the baselines (more substantial for the combination WN+WN-HypInfer). In most of the cases, however, the additional relations decrease the accuracy. For the WordNet relations, these improvement-inducing combinations include inference over the hypernymy relation (54.15) and inference over the second verb of the cause relation (49.25). For the Extended WordNet relations, one of the sets that outperforms the baseline includes inference over hypernymy, but the other one includes inference over antonymy.

5 Experiments with Semantic Relations in Extended WordNet

The Extended WordNet (Mihalcea and Moldovan, 2001) is constructed on the basis of analyses of the glosses of the synsets. During this analysis, the open class words were annotated with word

<i>KG</i>	<i>SemCor</i>
WN+WN-HypInfer	54.15
WN+WN-AntInfer	48.49
WN+WN-ClsInfer	48.48
WN+WN-Cs1stVerbInfer	49.21
WN+WN-Cs2ndVerbInfer	49.25
WN+WN-DerNAInfer	48.49
WN+WN-DerNNInfer	47.82
WN+WN-DerNVInfer	47.79
WN+WN-DerVNInfer	48.69
WN+WN-Ent1stVerbInfer	49.21
WN+WN-Ent2ndVerbInfer	49.21
WN+WN-InsInfer	48.89
WNG+WN-HypInfer	58.93
WNG+WN-AntInfer	59.08
WNG+WN-ClsInfer	57.66
WNG+WN-Cs1stVerbInfer	58.85
WNG+WN-Cs2ndVerbInfer	58.80
WNG+WN-DerNAInfer	58.41
WNG+WN-DerNNInfer	58.62
WNG+WN-DerNVInfer	55.68
WNG+WN-DerVNInfer	58.89
WNG+WN-Ent1stVerbInfer	58.84
WNG+WN-Ent2ndVerbInfer	58.79
WNG+WN-InsInfer	58.23

Table 3: Experimental results when using some of the inferred sets of relations. The results that are above the baselines from Table 1 are bolded.

synsets from PWN3.0. For example, the synset {stony coral, madrepore, madriporian coral} — 01916925-n, is defined by “corals having calcareous skeletons aggregations of which form reefs and islands.” After the analysis, the following synsets are selected: 02673969-a — *calcareous*, 01917882-n — *mushroom coral*, 05585383-n — *skeleton*, 07951464-n — *aggregation*, 09316454-n — *island*, 09406793-n — *reef*, and 02621395-v — *form*. Each of these synsets is related to the synset to which the gloss belongs to:

```

u:01916925-n v:02673969-a s:30g1c d:0
u:01916925-n v:01917882-n s:30g1c d:0
u:01916925-n v:05585383-n s:30g1c d:0
u:01916925-n v:07951464-n s:30g1c d:0
u:01916925-n v:09316454-n s:30g1c d:0
u:01916925-n v:09406793-n s:30g1c d:0
u:01916925-n v:02621395-v s:30g1c d:0

```

The first division of the relations in WNG into groups is on the basis of the parts of speech of the main synset. The four sets are: WNG-A (first synset is for adjectives), WNG-N (first synset is

for nouns), WNG-R (first synset is for adverbs), and WNG-V (first synset is for verbs).

<i>KG</i>	<i>SemCor</i>
WN+WNG-A	52.80
WN+WNG-N	56.85
WN+WNG-R	51.56
WN+WNG-V	52.61

Table 4: Experimental results when using the sets of relations from the XWN knowledge graph.

In Table 4 we present the impact of each of these sets of relations on the knowledge graph of WN. As can be seen, each set adds accuracy above the baseline of WN. When comparing the inferred relations (Table 3) and the WNG sets, it can be observed that the set WNG-N improves accuracy even over the WN-HypInfer set.

Additionally, each of the groups — WNG-A, WNG-N, WNG-R, and WNG-V — was divided into four subgroups on the basis of the part of speech of the second synset in the relation. Thus, we created 16 new sets: WNG-AA, WNG-AN, ..., WNG-VV. After experimenting with each of them, we arrived at the following combination: WN, WNG-AN, WNG-NN, WNG-RN, and WNG-VN. The accuracy for this combination is **56.99**, which is higher than the results for each individual set.

6 Syntax-based Relations

As was mentioned above, in our experiments we have also used semantic relations from a syntactically annotated corpus. To achieve this, we parsed SemCor with a dependency parser included in IXA pipeline. Then we divided the corpus in a proportion one-to-three: first part comprises of 49 documents (from br-a01 to br-f44) and it was used as a test set in the experiments reported here. The rest of the documents formed the training set from which the new relations were extracted. First, we defined patterns of dependency relations. For example, we used patterns like the following: $s_1 \text{subj} s_2$, which defines a relation between a noun synset s_1 and a verb synset s_2 ; $s_1 \text{mods} s_2$, which defines a relation between an adjective synset s_1 and a noun synset s_2 ; $s_1 \text{modxpobj} s_2$, which defines a relation between a noun synset s_1 and a noun synset s_2 ; etc. We extracted the following sets of relations: SC-AA, SC-AN, SC-AV, SC-NN, SC-NV, SC-RA, SC-RN, SC-RR, SC-RV,

SC-VN, SC-VV, where the suffixes — AA, AN, AV, etc. — denote the parts of speech of the related synsets. The results from the experiments performed are presented in Table 5. As can be seen, many of the extracted new sets increase the accuracy above the baseline for the original knowledge graph — WNG.

<i>KG</i>	<i>SemCor</i>
WNG+SC-AA	59.08
WNG+SC-AN	59.13
WNG+SC-AV	59.28
WNG+SC-NN	58.69
WNG+SC-NV	59.20
WNG+SC-RA	59.35
WNG+SC-RN	58.77
WNG+SC-RR	58.92
WNG+SC-RV	59.24
WNG+SC-VN	58.92
WNG+SC-VV	59.09

Table 5: Results from experiments using the sets of relations from syntax.

We have combined most of these sets in joint combinations. The combination of all the sets with the original knowledge graph: WNG, SC-AA, SC-AN, SC-AV, SC-NN, SC-NV, SC-RA, SC-RN, SC-RR, SC-RV, SC-VN, SC-VV gives accuracy of **60.13**. The best combination is WNG, SC-AA, SC-AN, SC-AV, SC-NV, SC-RA, SC-RN, SC-RR, SC-RV, SC-VN, SC-VV. The accuracy for this combination is **60.14**.

We also performed inference over these sets of relations using hypernymy and hyponymy hierarchies for nouns and verbs. The best result was achieved for the combination WNG, SC-AA, SC-AN, SC-AV, SC-NV, SC-RA, SC-RR, SC-RV, SC-VN, SC-VV, WN-HypInfer, WN-AntInfer. Its accuracy is **60.42**. This result is 1.5 % higher than the baseline for the original knowledge graph. This improvement is statistically significant.

7 Conclusion

In this paper we have evaluated the performance of different relations encoded in the knowledge graph, for the purposes of the knowledge-based Word Sense Disambiguation. Each of the sets of relations reflects an important linguistic piece of knowledge. Thus, each of them is important for the description of languages. However, from the

point of view of knowledge-based WSD each of these relations, as well as their various combinations, seem to have a different impact on the performance of the task.

The results from the experiments show that the addition of whole sets of relations might have a positive or a negative effect. In our view, at least two factors are of importance: (i) the number of relations assigned to each synset. Following Zipf's law, we can conclude that the distribution of relations per synset is very uneven. For many synsets there is not sufficient information present in the context, in order for a good decision to be taken. For many ambiguous words the context provides no information for disambiguation, and the decision is taken arbitrarily. (ii) The second factor is that the inference rules applied to the explicit relations do not produce the expected improvement. This might be due to the fact that WordNet is not the right place to store the inference information. Our expectations about the positive influence of inference are not always realized in practice. For instance, we expected to get relations between events and their participants from the derivational relations, but this was often not the case. If we take the verb "to kiss" and the derived noun "kisser", we would expect that "kisser" is a more general synset than the synsets for any specific kisser. But the synset for "kisser" had no single hyponym in WordNet. The gloss is *someone who kisses* and it determines the connection from "kisser" to "someone" who is the most general agent of the verb "to kiss". The connection is stated in XWN via the gloss of the noun "kisser". But for this configuration of relations in the original graph there is no inference rule defined. It seems that the systemic and monotonic knowledge that is needed for WSD and other NLP tasks is not always considered interesting enough to encode in various lexical resources.

Our future goals are the following: (i) the application of more complicated inference rules; (ii) the modification of relations per synset in order to ensure enough disambiguation relations. We expect these modifications of the relations to be performed via machine learning techniques over the contexts of the words in large corpora.

The number of synsets in the knowledge graph is 136,334, thus the possible links between them are in total 18,586,823,222. At the same time, the number of the actual links in the biggest graphs

used in practice, is less than 5 million, which is only 0.027 % of all possible combinations. Probably we need much more than 0.027 % of the links in order to capture all the available, and also all the necessary, knowledge for WSD. However, in such cases a faster algorithm must be employed.

Acknowledgements

This research has received support by the EC's FP7 (FP7/2007-2013) project under grant agreement number 610516: "QTLeap: Quality Translation by Deep Language Engineering Approaches."

References

- Eneko Agirre and Aitor Soroa. 2009. Personalizing PageRank for word sense disambiguation. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 33–41.
- Eneko Agirre, Oier López de Lacalle, and Aitor Soroa. 2014. Random walks for knowledge-based word sense disambiguation. *Comp. Ling.*, 40(1):57–84.
- Eneko Agirre, Ander Barrena, and Aitor Soroa. 2015. Studying the wikipedia hyperlink graph for relatedness and disambiguation. *arXiv preprint arXiv:1503.01655*.
- Sergey Brin and Lawrence Page. 2012. The anatomy of a large-scale hypertextual web search engine. *Computer networks*, 56(18):3825–3833.
- Christiane Fellbaum, editor. 1998. *WordNet An Electronic Lexical Database*. The MIT Press, Cambridge, MA ; London, May.
- Paweł Kdzia, Maciej Piasecki, Jan Koco, and Agnieszka Indyka-Piasecka. 2014. Distributionally extended network-based word sense disambiguation in semantic clustering of polish texts. *{IERI} Procedia*, 10:38 – 44. International Conference on Future Information Engineering (FIE 2014).
- Rada Mihalcea and Dan I. Moldovan. 2001. extended wordnet: progress report. In *in Proceedings of NAACL Workshop on WordNet and Other Lexical Resources*, pages 95–100.
- George A. Miller, Claudia Leacock, Randee Teng, and Ross T. Bunker. 1993. A semantic concordance. In *Proc. of HLT '93*, pages 303–308.
- Alexander Popov, Stanislava Kancheva, Svetlomira Manova, Ivaylo Radev, Kiril Simov, and Petya Osenova. 2014. The sense annotation of bultreebank. In *Proceedings of TLT13*, pages 127–136.
- Kiril Simov, Alexander Popov, and Petya Osenova. 2015. Improving word sense disambiguation with linguistic knowledge from a sense annotated treebank. In *Proc. of RANLP 2015.*, pages 596–603.