Grammatical Role Embeddings for Enhancements of Relation Density in the Princeton WordNet

Kiril Simov[1], Alexander Popov[1], Iliana Simova[2], Petya Osenova[1]
[1]IICT-BAS, Sofia, Bulgaria, {kivs|alex.popov|petya}@bultreebank.org
[2]Saarland University, Saarbrücken, Germany, ilianas@coli.uni-saarland.de

Abstract

In this paper we present an approach to train subatom embeddings for verbs. For each verb we learn not just one embedding, but several. One for the verb itself and embeddings for each grammatical role of this verb. For example, for the verb 'to give' we learn four embeddings: one for the lemma 'give', one for the subject, one for the direct object and one for the indirect object of it. We are exploiting these grammatical role embeddings in order to add new syntagmatic relations to WordNet. The evaluation of the quality of the new relations is done extrinsically via Knowledge-based Word Sense Disambiguation.

1 Introduction

In this paper we present an approach to extending the knowledge graph, based on Princeton English WordNet (PWN) — (Fellbaum, 1998) — with syntagmatic relations. Our aim is to improve the knowledge-based word sense disambiguation (KWSD). In several papers we showed that adding syntagmatic relations from syntactic and semantic annotated corpora improves the performance of KWSD — (Simov et al., 2015) and (Simov et al., 2016). The main types of syntagmatic relations extracted from these corpora are the ones corresponding to the grammatical roles: verb-subject (`has-subj`), verb-direct object (`has-dobj`) and verb-indirect object (`has-iobj`). Although we managed to extract good sets of new relations, the main problem is that corpora annotated with semantic and syntactic information contain only a fraction of all the possible syntagmatic relations.

The inheritance over the hierarchies of PWN is problematic because the hierarchies of PWN are not monotonic. For that reason, in this paper we use feature learning in low dimensional vectors of real numbers known as *embeddings*.

Word Embeddings play an important role in the new stream of natural language processing applications, providing latent features for lexical items. It is expected that the necessary features are encoded within the embedding space. For example, a verb embedding represents information for its valency frame elements. Unfortunately, we can check this information only indirectly. In the paper we report embeddings on the subatom level[1] that make explicit some of the features related to the semantic selectional restrictions on grammatical roles of words in text. Thus our goal is not to learn an embedding for a verb, but rather embeddings for the participants in the event (or state) denoted by that verb.

Such an explicit embedding of the valency frame elements has many potential applications. In this work we exploit these embeddings for adding new syntagmatic relations to PWN with the aim to improve applications such as KWSD. Evaluation in the paper is performed by automatically extending WordNet with ranked relations within the context of KWSD. We show that adding higher ranked relations improves the performance of KWSD. Further we provide manual inspection and validation of the new relations that also supports the feasibility of our approach.

Our method is similar to the other popular methods for relation extraction. The main difference is that we do not implement relation embeddings, but rather a general embedding for one of the entities involved in the rela-

---

[1]By subatom level we mean the arguments of a predicate.

tion. Also we work with relations that are not present in the knowledge source we extend — PWN in our case. In this way we hope that our method is applicable also to the under-resourced languages.

The structure of the paper is as follows: Section 2 briefly discusses related work. In section 3 we present our motivation to extend WordNet with syntagmatic relations. Section 4 outlines an example of subatom sentential semantics based on the ideas behind Minimal Recursion Semantics. Section 5 describes the mechanism for creating grammatical role embeddings. In section 6 the experiment setup is presented and the results are discussed. The last section concludes the paper.

2   Related Work

The success of KWSD approaches apparently depends on the quality of the knowledge graph – whether the knowledge represented in terms of nodes and relations (arcs) 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. With respect to the extension of WordNet with syntagmatic information there exist many works such as (Bentivogli and Pianta, 2004) and (Lothar Lemnitzer and Gupta, 2008).

Here we present in more detail only one approach similar to ours. It is described in Agirre and Martinez (2002) and explores the extraction of syntactically supported semantic relations from manually annotated corpora. In this line of research SemCor — (Miller et al., 1993), being a semantically annotated corpus, was processed with the MiniPar dependency parser and the subject-verb and object-verb relations were consequently extracted. The new relations were represented on several levels: as word-to-class and class-to-class relations. The extracted selectional relations were then added to WordNet and used in the WSD task. The main differences with the approach described here are as follows: we used a bigger set of relations (since it includes also indirect-object-to-verb relations). Apart from that, the new relations reported in this paper are not added as selectional relations, but as semantic

relations between the corresponding synsets. This means that the specific syntactic role of the participant is not taken into account, but only the connectedness between the participant and the event is registered in the knowledge graph. Also, in our work we use embeddings as filters, instead of the selectional restrictions approach undertaken in Agirre and Martinez (2002).

There is also a huge number of works on extending world knowledge oriented graphs with new relations (see (Minervini et al., 2015), and (Nguyen et al., 2016) among others). The main difference in our case is that we do not learn instances of the required relations from the corpora, but we learn semantic restrictions over the arguments of the relations. The candidate relations are generated from knowledge base itself (WordNet here).

3   WordNet Extensions with New Relations

As mentioned above, in our previous works we extended PWN with syntagmatic relations using semantically annotated corpora such as SemCor. The idea was that if there is a subject-verb syntactic relation in the corpus, and the related verb and noun are manually annotated with synset ids from PWN, we could reliably assume that there is a semantic relation between the noun and the verb synsets in PWN. At a more general level we call this relation `has-participant`. It is directed from the verb to the noun synset. In order to draw a distinction between the different participants in an event (state), we use subrelations named after their grammatical roles: `has-subj`, `has-dobj`, and `has-iobj`.[2]

Adding a `has-participant` relation between two synsets in WordNet imposes two questions: (1) Does this relation hold for more specific synsets? (2) Does this relation generalize to more general synsets? In our previous research on extending WordNet with new relations from semantically and syntactically annotated corpora — (Simov et al., 2015) and (Simov et al., 2016) — we showed that using inference over the WordNet hierarchy adds new appropriate relations between verb

---

[2]In future work we plan to switch to semantic role names.

synsets and noun synsets. Especially with respect to the `has-participant` relation, we assume that the relation holds when the noun synset is substituted with a hyponymic synset and that it also holds when the verb synset is substituted with a hypernymic synset. We noticed that in many cases such an inheritance is not correct. For example, if we have "A **doctor** *operates* a patient", it does not entail that all doctors can operate. Thus we cannot reliably substitute the synset for 'doctor' with each of its hyponymic synsets. It is also true that the verb synset allows many more participants than the instances in the corpus. For example, "A **surgeon** *cures* a patient" does not imply that only hyponymic synsets are appropriate to substitute 'surgeon'. Thus, although the extraction from syntactically and semantically annotated corpora is a reliable method for adding syntagmatic relations to WordNet, their generalization to all possible syntagmatic relations is problematic. Another problem is that such manually annotated corpora are relatively small and many verbs and nouns do not appear in them. Thus, we need a new mechanism for selection of appropriate noun synsets for participants of verbs. In this paper we used subatom semantic embeddings for checking which ones are appropriate. Such subatom semantic embeddings for each verbal synset are constructed for the appropriate grammatical roles: subject, direct object and indirect object. Having these embeddings, we rank each noun synset in PWN with respect to the corresponding grammatical role. The closer the noun synset embedding to the grammatical role embedding, the more appropriate is the noun synset as a participant for the corresponding grammatical roles in the selected verbal synset. In the rest of the paper we present some additional motivation why such subatomic embeddings are useful, how we could train and evaluate them.

4   Minimal Recursion Semantics

An additional piece of motivation for subatom semantic embeddings is the construction of a logical form for a sentence. In many semantic theories the lexical semantics is represented not only by using predicates from first order logic, but by exploring a more complicated schema which would allow access to a more detailed representation of the semantic interpretation. As an illustration of such a kind of semantics we assume Minimal Recursion Semantics (MRS) — (Copestake et al., 2005). An MRS structure is a tuple <GT, R, C>, where GT is the top handle, R is a bag of EPs (elementary predicates) and C is a bag of handle constraints, such that there is no handle h that outscopes GT. Each elementary predication contains exactly four components: (1) a handle which is the label of the EP; (2) a relation; (3) a list of zero or more ordinary variable arguments of the relation; and (4) a list of zero or more handles corresponding to scopal arguments of the relation (i.e., holes). Here is an example of an MRS structure for the sentence *"Every dog chases some white cat."*

<h0, {h1:every(x,h2,h3), h2:dog(x),
    h4:chase(e, x, y), h5:some(y,h6,h7),
    h6:white(y), h6:cat(y)}, {}>

The top handle is h0. The quantifiers are represented as the relations $every(x, y, z)$ and $some(x, y, z)$, where x is the bound variable, y and z are handles determining the restriction and the body of the quantifier. The conjunction of two or more relations is represented by sharing the same handle (h6 above). The outscope relation is defined as a transitive closure of the immediate outscope relation between two elementary predications — EP immediately outscopes EP' iff one of the scopal arguments of EP is the label of EP'. In the example the set of handle constraints is empty, which means that the representation is underspecified with respect to the scope of both quantifiers.

In order to use semantic embeddings over MRS structures we need to determine the interactions of the latent features for each of the predicate arguments. For example, the features from the embeddings for 'every', 'dog', and 'chase' have to agree on the common argument denoted by the variable 'x'. In order to control this interaction in a better way, we would like for each multiargument predicate to learn an embedding per argument. Thus for the above MRS structure we will need to have embeddings for 'x', 'y', 'e', 'h0', ... 'h7'. When we have them, we would like also to create an embedding related to the first argument

of 'every'. The argument of 'dog' and the second argument of 'chase' have to "agree".

Our long-term goal is to train such subatom embeddings. Here we present an approach for learning such embeddings for grammatical roles. Then we use these embeddings for extending of WordNet with syntagmatic relations, as it was described above.

5 Grammatical Role Embeddings from Parsed Corpora

In our first experiment we learned subatom semantic embeddings on the basis of dependency-parsed corpora. We determined the arguments as wordforms in the text. As an example, for the above mentioned case we used the position of 'dog'. In order to generalize over the different word forms in the different examples in the corpus we substituted the wordforms for the corresponding argument with a pseudoword form. For example, in the above sentence we generated the following variations with pseudoword forms for the different arguments of the different predicates:

Every *SUBJ_chase* chases some white cat.

Every dog chases some white *OBJ_chase.*

and many more. Having learned embeddings for these pseudowords, we assume that they represent the selectional features for the corresponding grammatical roles of the verbs.

The actual corpus we have used is WaCkypedia_EN corpus — (Baroni et al., 2009). The WaCkypedia_EN corpus was reparsed with a more recent version of the Stanford CoreNLP dependency parser. The dependency of type "collapsed-cc" was selected, which collapses several dependency relations in order to obtain direct dependencies between content words, and in addition propagates dependencies involving conjuncts. For instance, a parse of the sentence "the dog runs and barks" would result in the relations nsubj(dog, runs) and nsubj(dog, barks). This type of dependency allows for a token to have multiple head words.

The head word of each noun phrase subject, as well as direct and indirect object, is then replaced by its predicate role and its governing verb's lemma (SUBJ_run, SUBJ_bark — both for the noun 'dog'). When a token has more than one head word suitable for substi-

tution, copies of the sentence are created for each alternative replacement.

For the relation `has-subj` we use the dependency relations 'nsubj' and 'nsubjpass'; for the relation `has-dobj` we use the dependency relation 'dobj'; and for the relation `has-iobj` we use the dependency relation 'iobj'. In order to minimize some errors we enforced a condition that the dependency word should be a noun.

6 Experiments and Results

In this section we describe the experimental set up and the results.

Corpora preparation.

The corpora that the algorithms for word embeddings are trained on can contain either natural language text (e.g. Wikipedia or newswire articles) or artificially generated pseudo texts. Such pseudo texts can be the output from the Random Walk algorithm, when it is set to the mode of selecting sequences of nodes from a knowledge graph (KG) — see (Goikoetxea et al., 2015) for generation of pseudo corpora from a Word-Net knowledge graph and (Ristoski and Paulheim, 2016) for generation of pseudo corpora from RDF knowledge graphs such as DBPedia, GeoNames, FreeBase. Here we report results only for knowledge graphs based on WordNet and its extensions.

The corpus for training of the embeddings reported here consists of two parts: (1) pseudo corpus generated over WordNet (PCWN); and (2) real text corpora (RTC). PCWN is used to ensure that the embeddings represent features extracted from the knowledge within the WordNet. RTC is used to represent relevant contexts for learning embeddings of pseudo words for subjects, direct objects and indirect objects. As RTC we have used WaCkypedia_EN corpus processed as described in Section 5.

The union of both corpora is used in the experiments. In RTC all the words were substituted with their lemmas. Punctuation marks and numbers were deleted. The PCWN corpus first was generated on the level of synset ids, then for each synset a lemma was selected from the synset randomly. The resulting corpus consists of lemmas and pseudowords for the grammatical roles. We used the Word2Vec

tool[3] in order to train the embeddings. From the various models we select the one with the best score on the similarity task. This model was trained with the following settings: context window of 5 words; 7 iterations; negative examples set to 5; and frequency cut sampling set to 7. The resulting embedding is lemma and pseudoword embedding. Training on the joint corpus ensure that the noun embeddings and pseudoword embeddings are in the same vector space and thus they are comparable.

Since the synset embeddings are not directly available, we need to calculate those. Thus, for each synset, we obtain its vector by averaging the vectors for all lemmas it can be expressed with (this information is retrieved from WordNet). For grammatical roles, we average the corresponding grammatical role vectors per each lemma in the particular verb synset; in this way, if a particular synset comprises N lemmas, we will average the vectors for $SUBJ\_lemma_1$, $SUBJ\_lemma_2$, ..., $SUBJ\_lemma_N$.

The first experiments with these embeddings showed some, but very small, improvements for the task of Knowledge-based Word Sense Disambiguation. The explanation for these results is that calculating synset embeddings on the bases of lemma embeddings is not good enough because of the high level of ambiguity of lemmas in PWN.

This is why we performed two more experiments with two new versions of the corpora. First, we annotated the RTC with senses using UKB system[4] for knowledge-based word sense disambiguation. For the PCWN corpus we have used the version generated only using synset ids. In this case the embeddings are directly trained over synsets. Unfortunately, this approach did not improve the results significantly. Our explanation for this is the fact that the annotation with UKB, even with our best knowledge graph from (Simov et al., 2016), is under 68% accuracy. This result is too low for our task. Second, we used the POS annotation for RTC to substitute each word with lemma-POS strings. In this way we differentiated the same lemma used as different parts-of-speech. For PCWN it is straightforward to substitute the synset ids with the combinations lemma-POS. This experiment demonstrated the best results which we report here. From these corpora we trained two embeddings: (1) embedding trained over RTC only[5]. We denote this embedding as `RTC`; and (2) embedding trained over the joint corpus. We denote this embedding as `RTCPCWN`.

Selection and Ranking of Candidate Relations.

The candidate relations are selected in the following way. For each verbal synset that has at least one grammatical role embedding we form candidate relations in the following format:

`u:noun-synset-id v:verb-synset-id`

where noun-synset-id is any noun synset in PWN. Thus, for each verb we generate more that 74 000 candidate relations. Here is an example:

`u:00031264-n v:02005948-v`

for 'arrive' (02005948-v) and 'group' for (00031264-n).

After the completion of this step, we have all the information necessary to compare synset embeddings with grammatical role embeddings that match verb synsets. The comparison is carried out by calculating the cosine similarity measure. By setting a similarity threshold, the filter can be controlled, so that more or fewer new relations are added to the extended graph. The same procedure is repeated for DOBJ and IOBJ relations. Using this approach for each candidate relation we calculate the cosine similarity measure between the noun synset embedding and the embedding for the corresponding grammatical role. We then used the result as a rank over the candidate relations.

Experiments with Knowledge-based Word Sense Disambiguation.

In order to check the usefulness of the added relations, we performed experiments with the UKB system[6] for knowledge-based word sense disambiguation. The UKB tool requires two resource files to annotate the input text — a dictionary file with all lemmas that can be possibly linked to a sense identifier. In our case

---

[3] https://code.google.com/archive/p/word2vec/
[4] http://ixa2.si.ehu.es/ukb/

[5] This was suggested to us by one of the reviewers in order to see the impact of adding PCWN.
[6] http://ixa2.si.ehu.es/ukb/

Knowledge Graph	SemCor	M13 SemeVal
wn30	51.56	48.41
wn30RTC40	50.32	49.51
wn30RTC45	**52.60**	49.57
wn30RTC47	50.20	48.47
wn30RTC50	50.34	49.63
wn30RTC52	50.58	**51.88**
wn30RTC55	51.05	51.70
wn30RTC57	51.60	51.52

Table 1: Results about relations ranked by embeddings from POS tagged real text corpus. The improvement for SemCor is **1.04** and for M13 SemeVal is **3.47**.

Knowledge Graph	SemCor	M13 SemeVal
wn30	51.56	48.41
wn30RTCPCWN35	51.88	49.27
wn30RTCPCWN38	53.68	51.39
wn30RTCPCWN40	53.91	**51.45**
wn30RTCPCWN42	**54.33**	50.42
wn30RTCPCWN43	54.08	50.18
wn30RTCPCWN44	52.56	49.93

Table 2: Results about relations ranked by embeddings from POS tagged real text corpus and pseudo corpus. The improvement for SemCor is **2.77** and for M13 SemeVal is **3.04**.

WordNet-derived relations were used for our knowledge base; consequently, the sense identifiers are WordNet IDs. For instance, a line from the WordNet extracted dictionary looks like this:

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

First comes the lemma associated with the relevant word senses, after the lemma the sense identifiers are listed. 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 the associated word senses to the word form in the context and attempts to disambiguate its meaning among them.

The second resource file required for running the tool is the set of relations used to construct the knowledge graph over which UKB is run. The distribution of UKB comes with 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 KG 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.

In our experiments we relied on the following knowledge graphs: **wn30** — a knowledge graph formed from the relations in PWN (baseline); **wn30RTCNN** — a knowledge graph formed on the basis of **wn30** extended by the grammatical role-based relations, ranked by `RTC` embeddings. The number NN is the rank threshold for selection of the new relations. If NN is 47, then all relations with rank equal or higher than 47 are selected; **wn30RTCPCWNNN** — a knowledge graph formed on the basis of **wn30** extended by the

Role	Good	Acceptable	Bad
Subject	68	28	4
Direct object	67	24	9

Table 3: These are the manual evaluation results of the first 100 suggested relations selected via RTCPCWN embeddings for subject and direct object roles.

grammatical role-based relations, ranked by `RTCPCWN` embeddings. The interpretation of NN is the same.

The evaluation of the Word Sense Disambiguation is done over two test data sets: the test part of SemCor as defined in (Simov et al., 2015) and (Simov et al., 2016) and the English part of the test data set for the Multilingual Word Sense Disambiguation[7] — named here M13 SemeVal. The results are presented in Table 1 with improvement of **1.04** for SemCor and **3.47** for M13 SemeVal and in Table 2 with improvement of **2.77** for SemCor and **3.04** for M13 SemeVal. As it can be seen, the results depend on the type of the test corpus: SemCor is a balanced one and hence shows usages of many senses; M13 SemeVal is a smaller one and does not provide so many diverse types of text. All the results show that there is a rank for which there is a highest result, and for lower or higher ranks the result drops. Our explanation of this is that: (1) for higher ranks the number of added relations is smaller and thus their impact on the result is smaller; and (2) for the lower ranks the number of the not-so-good relations is higher. The impact of the PCWN embeddings is with respect to the type of the test corpora. In our view better relations are selected for a wider set of verbs.

The results show that the presented approach for selecting syntagmatic relations is feasible. Much more work is necessary to evaluate the grammatical role embedding.

6.1 Manual Inspection

The results were manually evaluated for the first 100 top-ranked subject and direct object relations. A scale was used that classifies the examples into the following groups: **good**, **acceptable**, and **bad**. The labels are correlative. This is possibly due to the fact that most verbs have intransitive and transitive usages.

As it can be seen from the table, most relations have been labeled as 'good', then come the 'acceptable' relations, and finally some 'bad' ones.

For both syntactic labels it was observed that the most frequent among the top-ranked relations are chemistry-oriented domain ones, such as: <dimethylglyoxime, dehydrogenate>. For the 'good' relation one example is as follows <streusel, caramelize>: "The streusel seeps down and caramelizes the apples in the most glorious way".

As acceptable relations we marked mostly ones that are good semantic relations but would not generate reasonable sentences because they are derivationally related. For example: <celebration, celebrate>, <chart, chart>, <oxidation, oxidate>, <measurement, measure>, etc. As bad example the following relation is considered: <cassareep, splinter>.

Thus manual evaluation also shows that the proposed mechanism of adding syntagmatic relations to PWN is feasible.

7 Conclusion

This paper presents an approach for learning features by subtomic semantic representation. It is useful for addition of syntagmatic relations to WordNet. Our longterm plan is to design a learning approach for each semantic argument of predicates in a logical form. The results here are the first steps in this direction.

In future we plan to do the following: (1) to include more arguments in the learning process like arguments of relational nouns and adjectives, which will impose mutual constraints on the learned features; (2) to experiment with different algorithms for learning of the embeddings such as (Levy and Goldberg, 2014), where it is possible to select arbitrary contexts which could be more appropriate for grammatical role embeddings learning; (3) to improve sense annotation in order to improve

---

[7] `https://www.cs.york.ac.uk/semeval-2013/task12/`

the sense embeddings; (4) to evaluate the sub-atom embeddings in other tasks such as coreference resolution, neural network word sense disambiguation; and (5) to perform tuning to the linguistic knowledge already represented in WordNet, FrameNet and other lexical resources and manually annotated corpora, by techniques similar to retrofitting.

Acknowledgements

References

Eneko Agirre and David Martinez. 2002. Integrating selectional preferences in wordnet. In *Proceedings of First International WordNet Conference*.

Marco Baroni, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetta. 2009. The wacky wide web: a collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, 43(3):209–226.

Luisa Bentivogli and Emanuele Pianta. 2004. Extending wordnet with syntagmatic information. In *In Second Global WordNet Conference*, pages 47–53.

Ann Copestake, Dan Flickinger, Carl Pollard, and Ivan Sag. 2005. Minimal recursion semantics: An introduction. *Research on Language & Computation*, 3(4):281–332.

Christiane Fellbaum, editor. 1998. *WordNet An Electronic Lexical Database*. The MIT Press, Cambridge, MA ; London, May.

Josu Goikoetxea, Aitor Soroa, and Eneko Agirre. 2015. Random walks and neural network language models on knowledge bases. In *HLT-NAACL*, pages 1434–1439. The Association for Computational Linguistics.

Omer Levy and Yoav Goldberg. 2014. Dependency-based word embeddings. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 302–308, Baltimore, Maryland, June. Association for Computational Linguistics.

Holger Wunsch Lothar Lemnitzer and Piklu Gupta. 2008. Enriching germanet with verb-noun relations - a case study of lexical acquisition. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may.

George A. Miller, Claudia Leacock, Randee Tengi, and Ross T. Bunker. 1993. A Semantic Concordance. In *Proceedings of the workshop on Human Language Technology*, HLT '93, pages 303–308, Stroudsburg, PA, USA. Association for Computational Linguistics.

Pasquale Minervini, Claudia d'Amato, Nicola Fanizzi, and Floriana Esposito. 2015. Efficient learning of entity and predicate embeddings for link prediction in knowledge graphs. In *Proceedings of the 11th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2015), 2015.*, pages 26–37.

Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu, and Mark Johnson. 2016. Stranse: a novel embedding model of entities and relationships in knowledge bases. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 460–466, San Diego, California, June. Association for Computational Linguistics.

Petar Ristoski and Heiko Paulheim, 2016. *RDF2Vec: RDF Graph Embeddings for Data Mining*, pages 498–514. Springer International Publishing, Cham.

Kiril Simov, Alexander Popov, and Petya Osenova. 2015. Improving word sense disambiguation with linguistic knowledge from a sense annotated treebank. In *Proceedings of the International Conference Recent Advances in Natural Language Processing*, pages 596–603, Hissar, Bulgaria, September. INCOMA Ltd. Shoumen, BULGARIA.

Kiril Simov, Petya Osenova, and Alexander Popov, 2016. pages 130–139. Springer International Publishing, Cham.