Interpreting Semantic Relations in Noun Compounds via Verb Semantics

Su Nam Kim[†] and Timothy Baldwin[†‡]
† Computer Science and Software Engineering
University of Melbourne, Victoria 3010 Australia

and

‡ NICTA Victoria Research Lab
University of Melbourne, Victoria 3010 Australia

`{snkim,tim}@csse.unimelb.edu.au`

Abstract

We propose a novel method for automatically interpreting compound nouns based on a predefined set of semantic relations. First we map verb tokens in sentential contexts to a fixed set of seed verbs using `WordNet::Similarity` and Moby's Thesaurus. We then match the sentences with semantic relations based on the semantics of the seed verbs and grammatical roles of the head noun and modifier. Based on the semantics of the matched sentences, we then build a classifier using `TiMBL`. The performance of our final system at interpreting NCs is 52.6%.

1 Introduction

The interpretation of **noun compounds** (hereafter, NCs) such as *apple pie* or *family car* is a well-established sub-task of language understanding. Conventionally, the NC interpretation task is defined in terms of unearthing the underspecified semantic relation between the head noun and modifier(s), e.g. *pie* and *apple* respectively in the case of *apple pie*.

NC interpretation has been studied in the context of applications including question-answering and machine translation (Moldovan et al., 2004; Cao and Li, 2002; Baldwin and Tanaka, 2004; Lauer, 1995). Recent work on the automatic/semi-automatic interpretation of NCs (e.g., Lapata (2002), Rosario and Marti (2001), Moldovan et al. (2004) and Kim and Baldwin (2005)) has made assumptions about the scope of semantic relations or restricted the domain of interpretation. This makes it difficult to gauge the general-purpose utility of the different methods. Our method avoids any such assumptions while outperforming previous methods.

In seminal work on NC interpretation, Finin (1980) tried to interpret NCs based on hand-coded rules. Vanderwende (1994) attempted the automatic interpretation of NCs using hand-written rules, with the obvious cost of manual intervention. Fan et al. (2003) estimated the knowledge required to interpret NCs and claimed that performance was closely tied to the volume of data acquired.

In more recent work, Barker and Szpakowicz (1998) used a semi-automatic method for NC interpretation in a fixed domain. Lapata (2002) developed a fully automatic method but focused on nominalizations, a proper subclass of NCs.[1] Rosario and Marti (2001) classified the nouns in medical texts by tagging hierarchical information using neural networks. Moldovan et al. (2004) used the word senses of nouns based on the domain or range of interpretation of an NC, leading to questions of scalability and portability to novel domains/NC types. Kim and Baldwin (2005) proposed a simplistic general-purpose method based on the lexical similarity of unseen NCs with training instances.

The aim of this paper is to develop an automatic method for interpreting NCs based on semantic relations. We interpret semantic relations relative to a fixed set of constructions involving the modifier and head noun and a set of **seed verbs** for each semantic relation: e.g. *(the) family <u>owns</u> (a) car* is taken as evidence for *family car* being an instance of the POSSESSOR relation. We then attempt to map all instances of the modifier and head noun as the heads of NPs in a transitive sentential context onto our set of constructions via lexical similarity over the verb, to arrive at an interpretation: e.g. we would hope to predict that *possess* is sufficiently similar to *own* that *(the) family <u>possesses</u> (a) car* would be recognised as support-

---

[1] With nominalizations, the head noun is deverbal, and in the case of Lapata (2002), nominalisations are assumed to be interpretable as the modifier being either the subject (e.g. *child behavior*) or object (e.g. *car lover*) of the base verb of the head noun.

ing evidence for the POSSESSOR relation. We use a supervised classifier to combine together the evidence contributed by individual sentential contexts of a given modifier–head noun combination, and arrive at a final interpretation for a given NC.

Mapping the actual verbs in sentences to appropriate seed verbs is obviously crucial to the success of our method. This is particularly important as there is no guarantee that we will find large numbers of modifier–head noun pairings in the sorts of sentential contexts required by our method, nor that we will find attested instances based on the seed verbs. Thus an error in mapping an attested verb to the seed verbs could result in a wrong interpretation or no classification at all. In this paper, we experiment with the use of WordNet (Fellbaum, 1998) and word clusters (based on Moby's Thesaurus) in mapping attested verbs to the seed verbs. We also make use of CoreLex in dealing with the semantic relation TIME and the RASP parser (Briscoe and Carroll, 2002) to determine the dependency structure of corpus data.

The data source for our set of NCs is binary NCs (i.e. NCs with a single modifier) from the Wall Street Journal component of the Penn Treebank. We deliberately choose to ignore NCs with multiple modifiers on the grounds that: (a) 88.4% of NC types in the Wall Street Journal component of the Penn Treebank and 90.6% of NC types in the British National Corpus are binary; and (b) we expect to be able to interpret NCs with multiple modifiers by decomposing them into binary NCs. Another simplifying assumption we make is to remove NCs incorporating proper nouns since: (a) the lexical resources we employ in this research do not contain them in large numbers; and (b) there is some doubt as to whether the set of semantic relations required to interpret NCs incorporating proper nouns is that same as that for common nouns.

The paper is structured as follows. Section 2 takes a brief look at the semantics of NCs and the basic idea behind the work. Section 3 details the set of NC semantic relations that is used in our research, Section 4 presents an extended discussion of our approach, Section 5 briefly explains the tools we use, Section 6.1 describes how we gather and process the data, Section 6.2 explains how we map the verbs to seed verbs, and Section 7 and Section 8 present the results and analysis of our approach. Finally we conclude our work in Section 9.

2 Motivation

The semantic relation in NCs is the relation between the head noun (denoted "H") and the modifier(s) (denoted "M"). We can find this semantic relation expressed in certain sentential constructions involving the head noun and modifier.

(1) *family car*
    CASE: *family **owns** the <u>car</u>.*
    FORM: H $\boxed{\text{own}}$ M
    RELATION: POSSESSOR

(2) *student protest*
    CASE: *protest is **performed** by <u>student</u>.*
    FORM: M is $\boxed{\text{performed}}$ by H
    RELATION: AGENT

In the examples above, the semantic relation (e.g. POSSESSOR) provides an interpretation of how the head noun and modifiers relate to each other, and the seed verb (e.g. *own*) provides a paraphrase evidencing that relation. For example, in the case of *family car*, the *family* is the POSSESSOR of the *car*, and in *student protest*, *student(s)* are the AGENT of the *protest*. Note that voice is important in aligning sentential contexts with semantic relations. For instance, *family car* can be represented as *car is owned by family* (passive) and *student protest* as *student performs protest* (active).

The exact nature of the sentential context varies with different synonyms of the seed verbs.

(3) *family car*
    CASE: *Synonym=have/possess/belong to*
    FORM: H $\boxed{\text{own}}$ M
    RELATION: POSSESSOR

(4) *student protest*
    CASE: *Synonym=act/execute/do*
    FORM: M is $\boxed{\text{performed}}$ by H
    RELATION: AGENT

The verb *own* in the POSSESSOR relation has the synonyms *have, possess* and *belong to*. In the context of *have* and *possess*, the form of relation would be same as the form with verb, *own*. However, in the context of *belong to*, *family car*

would mean that the *car belongs to family*. Thus, even when the voice of the verb is the same (*voice=active*), the grammatical role of the head noun and modifier can change.

In our approach we map the actual verbs in sentences containing the head noun and modifiers to seed verbs corresponding to the relation forms. The set of seed verbs contains verbs representative of each semantic relation form. We chose two sets of seed verbs of size 57 and 84, to examine how the coverage of actual verbs by seed verbs affects the performance of our method. Initially, we manually chose a set of 60 seed verbs. We then added synonyms from Moby's thesaurus for some of the 60 verbs. Finally, we filtered verbs from the two expanded sets, since these verbs occur very frequently in the corpus (as this might skew the results). The set of seed verbs {*have, own, possess, belong_to*} are in the set of 57 seed verbs, and {*acquire, grab, occupy*} are added to the set of 84 seed verbs; all correspond to the POSSESSOR relation.

For each relation, we generate a set of constructional templates associating a subset of seed verbs with appropriate grammatical relations for the head noun and modifier. Examples for POSSESSOR are:

$$S(\{have, own, possess\}^{\text{V}}, M^{\text{SUBJ}}, H^{\text{OBJ}}) \quad (5)$$

$$S(\{belong\_to\}^{\text{V}}, H^{\text{SUBJ}}, M^{\text{OBJ}}) \quad (6)$$

where V is the set of seed verbs, M is the modifier and H is the head noun.

Two relations which do not map readily onto seed verbs are TIME (e.g. *winter semester*) and EQUATIVE (e.g. *composer arranger*). Here, we rely on an independent set of contextual evidence, as outlined in Section 6.1.

Through matching actual verbs attested in corpus data onto seed verbs, we can match sentences with relations (see Section 6.2). Using this method we can identify the matching relation forms of semantic relations to decide the semantic relation for NCs.

3 Semantic Relations in Compound Nouns

While there has been wide recognition of the need for a system of semantic relations with which to classify NCs, there is still active debate as to what the composition of that set should be, or indeed
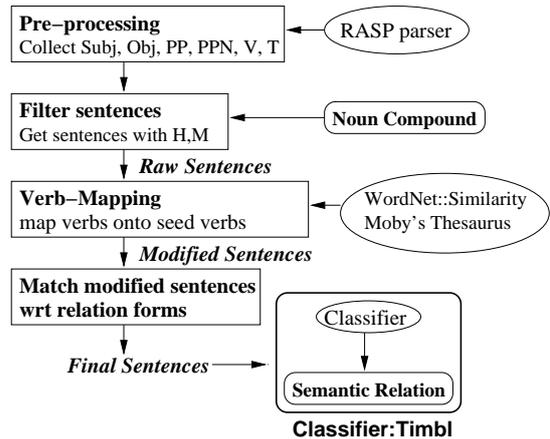

Figure 1: System Architecture

whether it is reasonable to expect that all NCs should be interpretable with a fixed set of semantic relations.

Based on the pioneering work on Levi (1979) and Finin (1980), there have been efforts in computational linguistics to arrive at largely task-specific sets of semantic relations, driven by the annotation of a representative sample of NCs from a given corpus type (Vanderwende, 1994; Barker and Szpakowicz, 1998; Rosario and Marti, 2001; Moldovan et al., 2004). In this paper, we use the set of 20 semantic relations defined by Barker and Szpakowicz (1998), rather than defining a new set of relations. The main reasons we chose this set are: (a) that it clearly distinguishes between the head noun and modifiers, and (b) there is clear documentation of each relation, which is vital for NC annotation effort. The one change we make to the original set of 20 semantic relations is to exclude the PROPERTY relation since it is too general and a more general form of several other relations including MATERIAL (e.g. *apple pie*).

4 Method

Figure 1 outlines the system architecture of our approach. We used three corpora: the Brown corpus (as contained in the Penn Treebank), the Wall Street Journal corpus (also taken from the Penn treebank), and the written component of the British National Corpus (BNC). We first parsed each of these corpora using RASP (Briscoe and Carroll, 2002), and identified for each verb token the voice, head nouns of the subject and object, and also, for each PP attached to that verb, the head preposition and head noun of the

NP (hereafter, **PPN**). Next, for our test NCs, we identified all verbs for which the modifier and head noun co-occur as subject, object, or PPN. We then mapped these verbs to seed verbs using `WordNet::Similarity` and Moby's Thesaurus (see Section 5 for details). Finally, we identified the corresponding relation for each seed verb and selected the best-fitting semantic relation using a classifier. To evaluate our approach, we built a classifier using `TiMBL` (Daelemans et al., 2004).

5 Resources

In this section, we outline the tools and resources employed in our method.

As our parser, we used RASP, generating a dependency representation for the most probable parse for each sentence. Note that RASP also lemmatises all words in a POS-sensitive manner.

To map actual verbs onto seed verbs, we experimented with two resources: `WordNet::Similarity` and Moby's thesaurus. `WordNet::Similarity`[2] is an open source software package that allows the user to measure the semantic similarity or relatedness between two words (Patwardhan et al., 2003). Of the many methods implemented in `WordNet::Similarity`, we report on results for one path-based method (WUP, Wu and Palmer (1994)), one content-information based method (JCN, Jiang and Conrath (1998)) and two semantic relatedness methods (LESK, Banerjee and Pedersen (2003), and VECTOR, (Patwardhan, 2003)). We also used a random similarity-generating method as a baseline (RANDOM).

The second semantic resource we use for verb-mapping method is Moby's thesaurus. Moby's thesaurus is based on Roget's thesaurus, and contains 30K root words, and 2.5M synonyms and related words. Since the direct synonyms of seed verbs have limited coverage over the set of sentences used in our experiment, we also experimented with using second-level indirect synonyms of seed verbs.

In order to deal with the TIME relation, we used CoreLex (Buitelaar, 1998). CoreLex is based on a unified approach to systematic polysemy and the semantic underspecification of nouns, and derives from WordNet 1.5. It contains 45 basic CoreLex types, systematic polysemous classes and 39,937 nouns with tags.

As mentioned earlier, we built our supervised classifier using `TiMBL`.

6 Data Collection

6.1 Data Processing

To test our method, we extracted 2,166 NC types from the Wall Street Journal (WSJ) component of the Penn Treebank. We additionally extracted sentences containing the head noun and modifier in pre-defined constructional contexts from the amalgam of: (1) the Brown Corpus subset contained in the Penn Treebank, (2) the WSJ portion of the Penn Treebank, and (3) the British National Corpus (BNC). Note that while these pre-defined constructional contexts are based on the contexts in which our seed verbs are predicted to correlate with a given semantic relation, we instances of all verbs occurring in those contexts. For example, based on the construction in Equation 5, we extract all instances of $S(V_i, M_j^{\text{SUBJ}}, H_j^{\text{OBJ}})$ for all verbs V_i and all instances of $NC_j = (M_j, H_j)$ in our dataset.

Two annotators tagged the 2,166 NC types independently at 52.3% inter-annotator agreement, and then met to discus all contentious annotations and arrive at a mutually-acceptable gold-standard annotation for each NC. The Brown, WSJ and BNC data was pre-parsed with RASP, and sentences were extracted which contained the head noun and modifier of one of our 2,166 NCs in subject or object position, or as (head) noun within the NP of an PP. After extracting these sentences, we counted the frequencies of the different modifier–head noun pairs, and filtered out: (a) all constructional contexts not involving a verb contained in WordNet 2.0, and (b) all NCs for which the modifier and head noun did not co-occur in at least five sentential contexts. This left us with a total of 453 NCs for training and testing. The combined total number of sentential contexts for our 453 NCs was 7,714, containing 1,165 distinct main verbs.

We next randomly split the NC data into 80% training data and 20% test data. The final number of test NCs is 88; the final number of training NCs varies depending on the verb-mapping method.

As noted in Section 2, the relations TIME and EQUATIVE are not associated with seed verbs. For TIME, rather than using contextual evidence, we simply flag the possibility of a TIME if the modifier is found to occur in the *TIME* class of CoreLex. In the case of TIME, we consider coordinated occur-
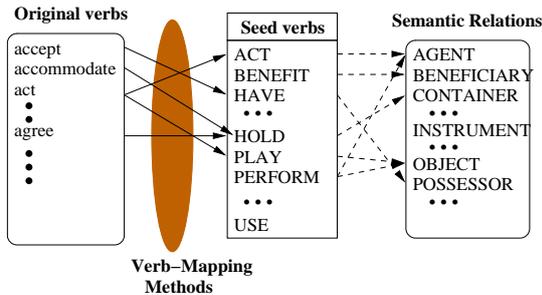
Figure 2: Verb mapping

Figure 3: Expanding synonyms

# of SeedVB	D-Synonyms	D/I-Synonyms
57	6,755(87.6%)	7,388(95.8%)
84	6,987(90.6%)	7,389(95.8%)

Table 1: Coverage of D and D/I-Synonyms

rences of the modifier and head noun (e.g. *coach and player* for *player coach*) as evidence for the relation.[3] We thus separately collate statistics from coordinated NPs for each NC, and from this compute a weight for each NC based on mutual information:

$$TIME(NC_i) = -log_2 \frac{freq(coord(M_i, H_i))}{freqM_i \times freq(H_i)} \quad (7)$$

where M_i and H_i are the modifier and head of NC_i, respectively, and $freq(coord(M_i, H_i))$ is the frequency of occurrence of M_i and H_i in coordinated NPs.

Finally, we calculate a normalised weight for each seed verb by determining the proportion of head verbs each seed verb occurs with.

6.2 Verb Mapping

The sentential contexts gathered from corpus data contain a wide range of verbs, not just the seed verbs. To map the verbs onto seed verbs, and hence estimate which semantic relation(s) each is a predictor of, we experimented with two different methods. First we used the WordNet::Similarity package to calculate the similarity between a given verb and each of the seed verbs, experimenting with the 5 methods mentioned in Section 5. Second, we used Moby's thesaurus to extract both direct synonyms (D-SYNONYM) and a combination of direct and second-level indirect synonyms of verbs (I-SYNONYM), and from this, calculate the closest-matching seed verb(s) for a given verb.

Figure 2 depicts the procedure for mapping verbs in constructional contexts onto the seed verbs. Verbs found in the various contexts in the corpus (on the left side of the figure) map onto one or more seed verbs, which in turn map onto one or more semantic relations.[4] We replace all non-seed verbs in the corpus data with the seed verb(s) they map onto, potentially increasing the number of corpus instances.

Since direct (i.e. level 1) synonyms from Moby's thesaurus are not sufficient to map all verbs onto seed verbs, we also include second-level (i.e. level 2) synonyms, expanding from direct synonyms. Table 1 shows the coverage of sentences for test NCs, in which **D** indicates direct synonyms and **I** indicates indirect synonyms.

7 Evaluation

We evaluated our method over both 17 semantic relations (without EQUATIVE and TIME) and the full 19 semantic relations, due to the low frequency and lack of verb-based constructional contexts for EQUATIVE and TIME, as indicated in Table 2. Note that the test data set is the same for both sets of semantic relations, but that the training data in the case of 17 semantic relations will not contain any instances for the EQUATIVE and TIME relations, meaning that all such test instances will be misclassified. The baseline for all verb mapping methods is a simple majority-class classifier, which leads to an accuracy of 42.4% for the TOPIC relation. In evaluation, we use two different values for our method: *Count* and *Weight*. *Count* is based on the raw number of corpus instances, while *Weight* employs the seed verb weight described in Section 6.1.

---

[3]Note the order of the modifier and head in coordinated NPs is considered to be irrelevant, i.e. *player and coach* and *coach and player* are equally evidence for an EQUATIVE interpretation for *player coach* (and *coach player*).
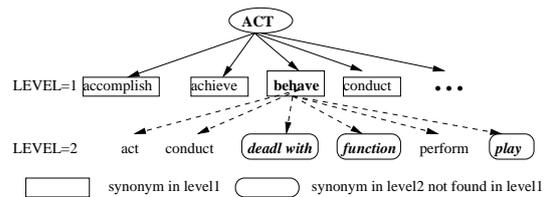
[4]There is only one instance of a seed verb mapping to multiple semantic relations, namely *perform* which corresponds to the two relations AGENT and OBJECT.

# of SR	# SeedV	Method	WUP	JCN	RANDOM	LESK	VECTOR	D-SYNONYM	I-SYNONYM
17		Baseline	.423	.423	.423	.423	.423	.423	.423
	57	Count	.324	.408	.379	.416	**.466**	.337	.337
		Weight	.320	.408	.371	.416	**.466**	.337	.342
	84	Count	.406	**.470**	.184	.430	.413	.317	.333
		Weight	.424	.426	.259	.457	**.526**	.341	.406
19		Baseline	.409	.409	.409	.409	.409	.409	.409
	57	Count	.315	.420	.384	.440	**.466**	.350	.337
		Weight	.311	.420	.376	.440	**.466**	.350	.342
	84	Count	.413	**.470**	.200	.414	.413	.321	.333
		Weight	.439	.446	.280	.486	**.526**	.356	.393

Table 2: Results with 17 relations and with 19 relations

#of SR	# SeedVB	Method	WUP	JCN	RANDOM	LESK	VECTOR
17		Baseline	.423	.423	.423	.423	.423
	57	Count	.423	.385	.379	.413	**.466**
		Weight	.423	.385	.379	.413	**.466**
	84	Count	.325	.439	.420	**.484**	.466
		Weight	.281	.393	.317	**.476**	.466
19		Baseline	.409	.409	.409	.409	.409
	57	Count	**.423**	.397	.392	.413	.413
		Weight	.423	.397	.392	.413	**.500**
	84	Count	.333	.439	.425	**.484**	.413
		Weight	.290	.410	.317	**.484**	.413

Table 3: Results of combining the proposed method and with the method of Kim and Baldwin (2005)

As noted above, we excluded all NCs for which we were unable to find at least 5 instances of the modifier and head noun in an appropriate sentential context. This exclusion reduced the original set of 2,166 NCs to only 453, meaning that the proposed method is unable to classify up to 80% of NCs. For real-world applications, a method which is only able to arrive at a classification for 20% of instances is clearly of limited utility, and we need some way of expanding the coverage of the proposed method. This is achieved by adapting the similarity method proposed by Kim and Baldwin (2005) to our task, wherein we use lexical similarity to identify the nearest-neighbour NC for a given NC, and classify the given NC according to the classification for the nearest-neighbour. The results for the combined method are presented in Table 3.

8 Discussion

For the basic method, as presented in Table 2, the classifier produced similar results over the 17 semantic relations to the 19 semantic relations. Using data from *Weight* and *Count* for both 17 and 19 semantic relations, the classifier achieved the best performance with VECTOR (context vector-based distributional similarity), followed by JCN and LESK. The main reason is that VECTOR is more conservative than the other methods at mapping (original) verbs onto seed verbs, i.e. the average number of seed verbs a given verb maps onto is small. For the other methods, the semantics of the original sentences are often not preserved under verb mapping, introducing noise to the classification task.

Comparing the two sets of semantic relations (17 vs. 19), the set with more semantic relations achieved slightly better performance in most cases. A detailed breakdown of the results revealed that TIME both has an above-average classification accuracy and is associated with a relatively large number of test NCs, while EQUATIVE has a below-average classification accuracy but is associated with relatively few instances.

While an increased number of seed verbs generates more training instances under verb mapping, it is imperative that the choice of seed verbs be made carefully so that they not introduce noise into the classifier and reducing overall performance. Figure 4 is an alternate representation of the numbers from Table 2, with results for each individual method over 57 and 84 seed verbs juxtaposed for each of *Count* and *Weight*. From this, we get the intriguing result that *Count* generally performs better over fewer seed verbs, while *Weight* performs better over more seed verbs.
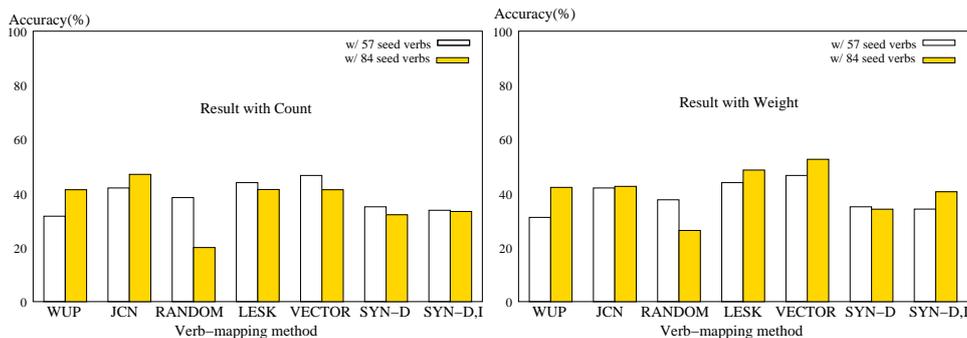
Figure 4: Performance with 57 vs. 84 seed verbs

#of SR	# SeedVB	WUP	LCH	JCN	LIN	RANDOM	LESK	VECTOR
17	Baseline	.433	.433	.441	.441	.433	.477	.428
	57	.449	.421	.415	.337	.409	.469	.344
	Baseline	.433	.433	.433	.433	.428	.438	.444
	84	**.476**	.416	.409	.349	.226	.465	.333
19	Baseline	.418	.418	.430	.430	.418	.477	.413
	57	.465	.418	.417	.341	.232	.462	.344
	Baseline	.413	.413	.418	.418	.413	.438	.426
	84	**.471**	.413	.407	.348	.218	.465	.320

Table 4: Results for the method of Kim and Baldwin (2005) over the test set used in this research

For the experiment where we combine our method with that of Kim and Baldwin (2005), as presented in Table 3, we find a similar pattern of results to the proposed method. Namely, VECTOR and LESK achieve the best performance, with minor variations in the absolute performance relative to the original method but the best results for each relation set actually dropping marginally over the original method. This drop is not surprising when we consider that we use an imperfect method to identify the nearest neighbour for an NC for which we are unable to find corpus instances in sufficient numbers, and then a second imperfect method to classify the instance.

Compared to previous work, our method produces reasonably stable performance when operated over the open-domain data with small amounts of training data. Rosario and Marti (2001) achieved about 60% using a neural network in a closed domain, Moldovan et al. (2004) achieved 43% using word sense disambiguation of the head noun and modifier over open domain data, and Kim and Baldwin (2005) produced 53% using lexical similarities of the head noun and modifier (using the same relation set, but evaluated over a different dataset). The best result achieved by our system was 52.6% over open-domain data, using a general-purpose relation set.

To get a better understanding of how our method compares with that of Kim and Baldwin (2005), we evaluated the method of Kim and Baldwin (2005) over the same data set as used in this research, the results of which are presented in Table 4. The relative results for the different similarity metrics mirror those reported in Kim and Baldwin (2005). WUP produced the best performance at 47-48% for the two relation sets, significantly below the accuracy of our method at 53.3%. Perhaps more encouraging is the result that the combined method—where we classify attested instances according to the proposed method, and classify unattested instances according to the nearest-neighbour method of Kim and Baldwin (2005) and the classifications from the proposed method—outperforms the method of Kim and Baldwin (2005). That is, the combined method has the coverage of the method of Kim and Baldwin (2005), but inherits the higher accuracy of the method proposed herein. Having said this, the performance of the Kim and Baldwin (2005) method over PRODUCT, TOPIC, LOCATION and SOURCE is superior to that of our method. In this sense, we believe that alternate methods of hybridisation may lead to even better results.

Finally, we wish to point out that the method as presented here is still relatively immature, with considerable scope for improvement. In its current form, we do not weight the different seed verbs

based on their relative similarity to the original verb. We also use the same weight and frequency for each seed verb relative to a given relation, despite seed verbs being more indicative of a given relation and also potentially occurring more often in the corpus. For instance, *possess* is more related to POSSESSOR than *occupy*. Also *possess* occurs more often in the corpus than *belong_to*. As future work, we intend to investigate whether allowances for these considerations can improve the performance of our method.

9 Conclusion

In this paper, we proposed a method for automatically interpreting noun compounds based on seed verbs indicative of each semantic relation. For a given modifier and head noun, our method extracted corpus instances of the two nouns in a range of constructional contexts, and then mapped the original verbs onto seed verbs based on lexical similarity derived from `WordNet::Similarity`, and Moby's Thesaurus. These instances were then fed into the `TiMBL` learner to build a classifier. The best-performing method was VECTOR, which is a context vector distributional similarity method. We also experimented with varying numbers of seed verbs, and found that generally the more seed verbs, the better the performance. Overall, the best-performing system achieved an accuracy of 52.6% with 84 seed verbs and the VECTOR verb-mapping method.

Acknowledgements

References

Timothy Baldwin and Takaaki Tanaka. 2004. Translation by machine of compound nominals: Getting it right. In *In Proceedings of the ACL 2004 Workshop on Multiword Expressions: Integrating Processing*, pages 24–31, Barcelona, Spain.

Satanjeev Banerjee and Ted Pedersen. 2003. Extended gloss overlaps as a measure of semantic relatedness. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, pages 805–810, Acapulco, Mexico.

Ken Barker and Stan Szpakowicz. 1998. Semi-automatic recognition of noun modifier relationships. In *Proceedings of the 17th international conference on Computational linguistics*, pages 96–102, Quebec, Canada.

Ted Briscoe and John Carroll. 2002. Robust accurate statistical annotation of general text. In *Proc. of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*, pages 1499–1504, Las Palmas, Canary Islands.

Paul Buitelaar. 1998. *CoreLex: Systematic Polysemy and Underspecification*. Ph.D. thesis, Brandeis University, USA.

Yunbo Cao and Hang Li. 2002. Base noun phrase translation using web data and the EM algorithm. In *19th International Conference on Computational Linguistics*, Taipei, Taiwan.

Walter Daelemans, Jakub Zavrel, Ko van der Sloot, and Antal van den Bosch. 2004. *TiMBL: Tilburg Memory Based Learner, version 5.1, Reference Guide*. ILK Technical Report 04-02.

James Fan, Ken Barker, and Bruce W. Porter. 2003. The knowledge required to interpret noun compounds. In *7th International Joint Conference on Artificial Intelligence*, pages 1483–1485, Acapulco, Mexico.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, USA.

Timothy W. Finin. 1980. *The semantic interpretation of compound nominals*. Ph.D. thesis, University of Illinois, Urbana, Illinois, USA.

Jay Jiang and David Conrath. 1998. Semantic similarity based on corpus statistics and lexical taxonomy. In *Proceedings on International Conference on Research in Computational Linguistics*, pages 19–33.

Su Nam Kim and Timothy Baldwin. 2005. Automatic interpretation of noun compounds using WordNet similarity. In *Second International Joint Conference On Natural Language Processing*, pages 945–956, JeJu, Korea.

Maria Lapata. 2002. The disambiguation of nominalizations. *Computational Linguistics*, 28(3):357–388.

Mark Lauer. 1995. *Designing Statistical Language Learners: Experiments on Noun Compounds*. Ph.D. thesis, Macquarie University.

Judith Levi. 1979. *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Dan Moldovan, Adriana Badulescu, Marta Tatu, Daniel Antohe, and Roxana Girju. 2004. Models for the semantic classification of noun phrases. In *HLT-NAACL 2004: Workshop on Computational Lexical Semantics*, pages 60–67, Boston, USA.

Siddharth Patwardhan, Satanjeev Banerjee, and Ted Pedersen. 2003. Using measures of semantic relatedness for word sense disambiguation. In *Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics*.

Siddharth Patwardhan. 2003. Incorporating dictionary and corpus information into a context vector measure of semantic relatedness. Master's thesis, University of Minnesota, USA.

Barbara Rosario and Hearst Marti. 2001. Classifying the semantic relations in noun compounds via a domain-specific lexical hierarchy. In *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing*, pages 82–90.

Lucy Vanderwende. 1994. Algorithm for automatic interpretation of noun sequences. In *Proceedings of the 15th Conference on Computational linguistics*, pages 782–788.

Zhibiao Wu and Martha Palmer. 1994. Verb semantics and lexical selection. In *32nd Annual Meeting of the Association for Computational Linguistics*, pages 133 –138.