

Automatic Interpretation of Noun Compounds Using WordNet Similarity

Su Nam Kim^{1,2} and Timothy Baldwin^{2,3}

¹ Computer Science, University of Illinois, Chicago, IL 60607 USA
sunamkim@gmail.com

² Computer Science and Software Engineering,
University of Melbourne, Victoria 3010 Australia

³ NICTA Victoria Lab, University of Melbourne, Victoria 3010 Australia
tim@csse.unimelb.edu.au

Abstract. The paper introduces a method for interpreting novel noun compounds with semantic relations. The method is built around word similarity with pre-tagged noun compounds, based on `WordNet::Similarity`. Over 1,088 training instances and 1,081 test instances from the Wall Street Journal in the Penn Treebank, the proposed method was able to correctly classify 53.3% of the test noun compounds. We also investigated the relative contribution of the modifier and the head noun in noun compounds of different semantic types.

1 Introduction

A **noun compound** (NC) is an \bar{N} made up of two or more nouns, such as *golf club* or *paper submission*; we will refer to the rightmost noun as the **head noun** and the remainder of nouns in the NC as **modifiers**. The interpretation of noun compounds is a well-researched area in natural language processing, and has been applied in applications such as question answering and machine translation [1,2,3]. Three basic properties make the interpretation of NCs difficult [4]: (1) the compounding process is extremely productive; (2) the semantic relationship between head noun and modifier in the noun compounds is implicit; and (3) the interpretation can be influenced by contextual and pragmatic factors.

In this paper, we are interested in recognizing the semantic relationship between the head noun and modifier(s) of noun compounds. We introduce a method based on word similarity between the component nouns in an unseen test instance NC and annotated training instance NCs. Due to its simplicity, our method is able to interpret NCs with significantly reduced cost. We also investigate the relative contribution of the head noun and modifier in determining the semantic relation.

For the purposes of this paper, we focus exclusively on binary NCs, that is NCs made up of two nouns. This is partly an empirical decision, in that the majority of NCs occurring in unrestricted text are binary,¹ and also partly due to there being existing methods for disambiguating the syntactic structure of higher-arity NCs, effectively decomposing them into multiple binary NCs [3]. Note also that in this paper, we

¹ We estimate that 88.4% of NCs in the Wall Street Journal section of the Penn Treebank and 90.6% of NCs in the British National Corpus are binary.

distinguish semantic relations from semantic roles. The semantic relation in an NC is the underlying relation between the head noun and its modifier, whereas its semantic role is an indication of its relation to the governing verb and other constituents in the sentence context.

There is a significant body of closely-related research on interpreting semantic relations in NCs which relies on hand-written rules. [5] examined the problem of interpretation of NCs and constructed a set of hand-written rules. [6] automatically extracted semantic information from an on-line dictionary and manipulated a set of hand-written rules to assign weights to semantic relations. Recently, there has been work on the automatic (or semi-automatic) interpretation of NCs [4,7,8]. However, most of this work is based on a simplifying assumption as to the scope of semantic relations or the domain of interpretation, making it difficult to compare the performance of NC interpretation in a broader context.

In the remainder of the paper, we detail the motivation for our work (Section 2), introduce the `WordNet::Similarity` system which we use to calculate word similarity (Section 3), outline the set of semantic relations used (Section 4), detail how we collected the data (Section 5), introduce the proposed method (Section 6), and describe experimental results (Section 7).

2 Motivation

Most work related to interpreting NCs depends on hand-coded rules [5]. The first attempt at automatic interpretation by [6] showed that it was possible to successfully interpret NCs. However, the system involved costly hand-written rules involving manual intervention. [9] estimated the amount of world knowledge required to interpret NCs and claimed that the high cost of data acquisition offsets the benefits of automatic interpretation of NCs.

Recent work [4,7,8] has investigated methods for interpreting NCs automatically with minimal human effort. [10] introduced a semi-automatic method for recognizing noun-modifier relations. [4] examined nominalizations (a proper subset of NCs) in terms of whether the modifier is a subject or object of the verb the head noun is derived from (e.g. *language understanding = understand language*). [7] assigned hierarchical tags to nouns in medical texts and classified them according to their semantic relations using neural networks. [8] used the word senses of nouns to classify the semantic relations of NCs. However, in all this work, there has been some underlying simplifying assumption, in terms of the domain or range of interpretations an NC can occur with, leading to questions of scalability and portability to novel domains/NC types.

In this paper, we introduce a method which uses word similarity based on WordNet. Word similarity has been used previously in various lexical semantic tasks, including word sense disambiguation [11,12]. [11] showed that term-to-term similarity in a context space can be used to disambiguate word senses. [12] measured the relatedness of concepts using similarity based on WordNet. [13] examined the task of disambiguating noun groupings with respect to word senses using similarity between nouns in NCs. Our research uses similarities between nouns in the training and test data to interpret the semantic relations of novel NCs.

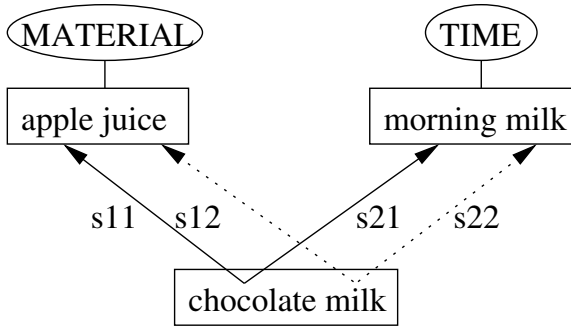


Fig. 1. Similarity between test NC *chocolate milk* and training NCs *apple juice* and *morning milk*

Table 1. WordNet-based similarities for component nouns in the training and test data

	Training noun	Test noun	S_{ij}
t1	apple	chocolate	0.71
t2	juice	milk	0.83
t1	morning	chocolate	0.27
t2	milk	milk	1.00

Figure 1 shows the correspondences between two training NCs, *apple juice* and *morning milk*, and a test NC, *chocolate milk*; Table 1 lists the noun pairings and noun–noun similarities based on WordNet. Each *training* noun is a component noun from the training data, each *test* noun is a component noun in the input, and S_{ij} provides a measure of the noun–noun similarity in *training* and *test*, where $t1$ is the modifier and $t2$ is the head noun in the NC in question. The similarities in Table 1 were computed by the WUP method [14] as implemented in `WordNet::Similarity` (see Section 3).

The simple product of the individual similarities (of each modifier and head noun, respectively) gives the similarity of the NC pairing. For example, the similarity between *chocolate milk* and *apple juice* is 0.60, while that between *chocolate milk* and *morning milk* is 0.27. Note that although *milk* in the input NC also occurs in a training exemplar, the semantic relations for the individual NCs differ. That is, while *apple juice* is *juice* made from *apples* (MATERIAL), *morning milk* is *milk* served in the *morning* (TIME). By comparing the similarity of both elements of the input NC, we are able to arrive at the conclusion that *chocolate milk* is more closely related to *chocolate milk*, which provides the correct semantic relation of MATERIAL (i.e. *milk* made from/flavored with *chocolate*). Unlike word sense disambiguation systems, our method does not need to determine the particular sense in which each noun is used. The next example (Table 2) shows how our method interprets NCs containing ambiguous nouns correctly.

One potential pitfall when dealing with WordNet is the high level of polysemy for many lexemes. We analyze the effects of polysemy with respect to *interest*. Assume that we have the two NCs *personal interest* (POSSESSION) and *bank interest* (CAUSE/TOPIC) in the training data. Both contain the noun *interest*, with the meaning of a state of cu-

Table 2. The effects of polysemy on the similarities between nouns in the training and test data

	Training noun	Test noun	S_{ij}
t1	personal	loan	0.32
t2	interest	rate	0.84
t1	bank	loan	0.75
t2	interest	rate	0.84

Table 3. Varying contribution of the head noun and modifier in predicting the semantic relation

Relative contribution of modifier/head noun	Relation	Example
modifier < head noun	PROPERTY	<i>elephant seal</i>
modifier = head noun	EQUATIVE	<i>composer arranger</i>
modifier > head noun	TIME	<i>morning class</i>

riosity or concern about something in *personal interest*, and an excess or bonus beyond what is expected or due in *bank interest*. Given the test NC *loan rate*, we would get the desired result of *bank interest* being the training instance of highest similarity, leading to *loan rate* being classified with the semantic relation of CAUSE/TOPIC. The similarity between the head nouns *interest* and *rate* for each pairing of training and test NC is identical, as the proposed method makes no attempt to disambiguate the sense of a noun in each NC context, and instead aggregates the overall word-to-word similarity across the different sense pairings. The determining factor is therefore the similarity between the different modifier pairings, and the fact that *bank* is more similar to *loan* than is the case for *personal*.

We also investigate the weight of the head noun and the modifier in determining overall similarity. We expect for different relations, the weight of the head noun and the modifier will be different. In the relation EQUATIVE, e.g., we would expect the significance of the head noun to be the same as that of the modifier. In relations such as PROPERTY, on the other hand, we would expect the head noun to play a more important role than the modifier. Conversely, with relations such as TIME, we would expect the modifier to be more important, as detailed in Table 3.

3 WordNet::Similarity

WordNet::Similarity² [12] is an open source software package developed at the University of Minnesota. It allows the user to measure the semantic similarity or relatedness between a pair of concepts (or word senses), and by extension, between a pair of words. The system provides six measures of similarity and three measures of relatedness based on the WordNet lexical database [15]. The measures of similarity are based on analysis of the WordNet isa hierarchy.

² www.d.umn.edu/~tpederse/similarity.html

The measures of similarity are divided into two groups: path-based and information content-based. We chose four of the similarity measures in `WordNet::Similarity` for our experiments: `WUP` and `LCH` as path-based similarity measures, and `JCN` and `LIN` as information content-based similarity measures. `LCH` finds the shortest path between nouns [16]; `WUP` finds the path length to the root node from the least common subsumer (LCS) of the two word senses that is the most specific word sense they share as an ancestor [14]; `JCN` subtracts the information content of the LCS from the sum [17]; and `LIN` scales the information content of the LCS relative to the sum [18].

In `WordNet::Similarity`, relatedness goes beyond concepts being similar to each other. That is, `WordNet` provides additional (non-hierarchical) relations such as `has-part` and `made-of`. It supports our idea of interpretation of NCs by similarity. However, as [19] point out, information on relatedness has not been developed as actively as conceptual similarity. Besides, the speed of simulating these relatedness effects is too slow to use in practice. Hence, we did not use any of the relatedness measures in this paper.

4 Semantic Relations

A semantic relation in the context of NC interpretation is the relation between the modifier and the head noun. For instance, *family car* relates to `POSSESSION` whereas *sports car* relates to `PURPOSE`. [20] defined complex nominals as expressions that have a head noun preceded by one or more modifying nouns or denominal adjectives, and offered nine semantic labels after removing opaque compounds and adding nominal non-predicating adjectives. [5] produced a diverse set of NC interpretations. Other researchers have identified alternate sets of semantic relations, or conversely cast doubts on the possibility of devising an all-purpose system of NC interpretations [21]. For our work, we do not intend to create a new set of semantic relations. Based on our data, we chose a pre-existing set of semantic relations that had previously been used for automatic (or semi-automatic) NC interpretation, namely the 20-member classification of [10] (see Appendix). Other notable classifications include that of [6] which contains 13 relations based on WH questions, making it ideally suited to question answering applications. However, some relations such as `TOPIC` are absent. [7] proposed 38 relations for the medical domain. Such relations are too highly specialized to this domain, and not suitable for more general applications. [8] defined 35 semantic relations for complex nominals and adjective phrases.

5 Data Collection

We retrieved binary NCs from the Wall Street Journal component of the Penn treebank. We excluded proper nouns since `WordNet` does not contain even high-frequency proper nouns such as *Honda*. We also excluded binary NCs that are part of larger NCs. In tagging the semantic relations of noun compounds, we hired two annotators: two computer science Ph.D students. In many cases, even human annotators disagree on the tag allocation. For NCs containing more than one semantic relation, the annotators were

judged to have agreed is there was overlap in at least one of the relations specified by them for a given NC. The initial agreement for the two annotators was 52.31%. From the disagreement of tagged relations, we observed that decisions between SOURCE and CAUSE, PURPOSE and TOPIC, and OBJECT and TOPIC frequently have lower agreement. For the NCs where there was no agreement, the annotators decided on a set of relations through consultation. The distribution of semantic relations is shown in the Appendix. Overall, we used 1,088 NCs for the training data and 1,081 NCs for the test data.

6 Method

Figure 2 shows how to compute the similarity between the i^{th} NC in the test data and j^{th} NC in the training data. We calculate similarities for the component nouns of the i^{th} NC in the test data with all NCs in the training data. As a result, the modifier and head noun in the i^{th} test NC are each associated with a total of m similarities, where m is the number of NCs in the training data. The second step is to multiply the similarities of the modifier and head noun for all NCs in the training data; we experiment with two methods for calculating the combined similarity. The third step is to choose the NC in the training data which is most similar to the test instance, and tag the test instance according to the semantic relation associated with that training instance.

Formally, S_A is the similarity between NCs $(N_{i,1}, N_{i,2})$ and $(B_{j,1}, B_{j,2})$:

$$S_A((N_{i,1}, N_{i,2}), (B_{j,1}, B_{j,2})) = \frac{((\alpha S1 + S1) \times ((1 - \alpha)S2 + S2))}{2} \tag{1}$$

where $S1$ is the modifier similarity (i.e. $S(N_{i,1}, B_{j,1})$) and $S2$ is head noun similarity (i.e. $S(N_{i,2}, B_{j,2})$); $\alpha \in [0, 1]$ is a weighting factor.

S_B is an analogous similarity function, based on the F-score:

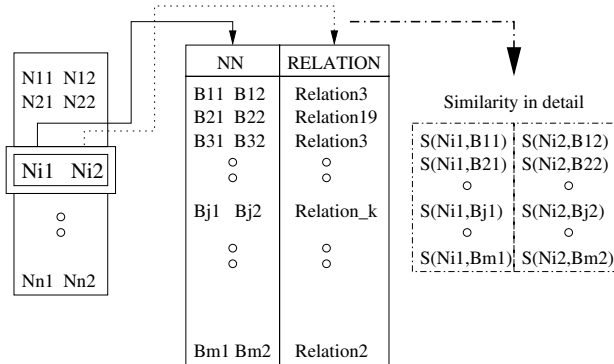


Fig. 2. Similarity between the i^{th} NC in the test data and j^{th} NC in the training data

$$S_B((N_{i,1}, N_{i,2}), B_{(j,1), B_{j,2}})) = 2 \times \frac{(S1 + \alpha S1) \times (S2 + (1 - \alpha)S2)}{(S1 + \alpha S1) + (S2 + (1 - \alpha)S2)} \quad (2)$$

The semantic relation is determined by *rel*:

$$\begin{aligned} rel(N_{i,1}, N_{i,2}) &= rel(B_{m,1}, B_{m,2}) \\ \text{where } m &= \underset{j}{\operatorname{argmax}} S((N_{i,1}, N_{i,2}), (B_{j,1}, B_{j,2})) \end{aligned} \quad (3)$$

7 Experimental Results

7.1 Automatic Tagging Using Similarity

In our first experiment, we tag the test NCs with semantic relations using four different measures of noun similarity, assuming for the time being that the contribution of the modifier and head noun is equal (i.e. $\alpha = 0.5$). The baseline for this experiment is a majority-class classifier, in which all NCs are tagged according to the TOPIC class.

Table 4. Accuracy of NC interpretation for the different WordNet-based similarity measures

Basis	Method	S_A	S_B
majority class	Baseline	465 (43.0%)	465 (43.0%)
path-based	WUP	576 (53.3%)	557 (51.5%)
path-based	LCH	572 (52.9%)	565 (52.3%)
information content-based	JCN	505 (46.7%)	470 (43.5%)
information content-based	LIN	512 (47.4%)	455 (42.1%)
human annotation	Inter-annotator agreement	565 (52.3%)	565 (52.3%)

Table 4 shows that **WUP**, using the S_A multiplicative method of combination, provides the highest NC interpretation accuracy, significantly above the majority-class baseline. It is particularly encouraging to see that **WUP** performs at or above the level of inter-annotator agreement (52.3%), which could be construed as a theoretical upper bound for the task as defined here. Using the F-score measure of similarity, **LCH** has nearly the same performance as **WUP**. Among the four measures of similarity used in this first experiment, the path-based similarity measures have higher performance than the information content-based methods over both similarity combination methods.

Compared to prior work on the automatic interpretation of NCs, our method achieves relatively good results. [7] achieved about 60% performance over the medical domain. [8] used a word sense disambiguation system to achieve around 43% accuracy interpreting NCs in the open domain. Our accuracy of 53% compares favourably to both of these sets of results, given that we are operating over open domain data.

7.2 Relative Contribution of Modifier and Head Noun

In the second experiment, we investigated the relative impact of the modifier and head noun in determining the overall similarity of the NC. While tagging the NCs, we got

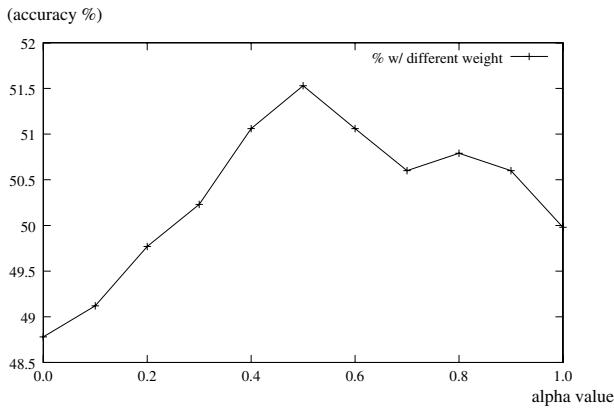


Fig. 3. Classifier accuracy at different α values

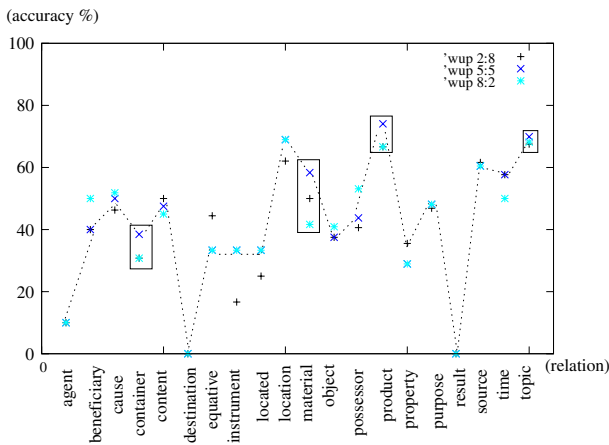


Fig. 4. Classification accuracy for each semantic relation at different α values

a sense of modifiers and head nouns having variable impact on the determination of the overall NC semantic relation. For this test, we used the WUP method based on our results from above and also because it operates over the scale [0, 1], removing any need for normalization. In this experiment, modifiers and head nouns were assigned weights (α in Equations 1 and 2) in the range 0.0, 0.1, ...1.0.

Figure 3 shows the relative contribution of the modifier and head noun in the overall NC interpretation process. Interestingly, the head noun seems to be a more reliable predictor of the overall NC interpretation than the modifier, and yet the best accuracy is achieved when each noun makes an equal contribution to the overall interpretation (i.e. $\alpha = 0.5$). Thus suggests that, despite any localized biases for individual NC interpretation types, the modifier and head noun have an equal impact on NC interpretation overall.

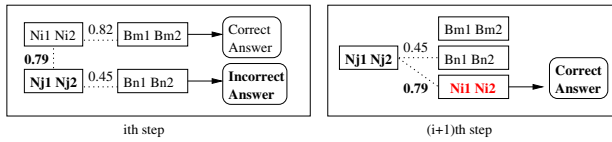


Fig. 5. Accumulating correctly tagged data

Figure 4 shows a breakdown of accuracy across the different semantic relation types for different weights. In Figure 4, we have shown only the weights 0.2, 0.5 and 0.8 (to show the general effect of variation in α). The dashed line shows the performance when the weight of modifiers and head nouns is the same ($\alpha = 0.5$). The \times symbol shows the results of modifier-biased interpretation ($\alpha = 0.8$) and the $+$ symbol shows the results of head noun-biased interpretation ($\alpha = 0.2$). From Figure 4, we can see that for relations such as CAUSE and INSTRUMENT, the modifier plays a more important role in the determination of the semantic relation of the NC. On the other hand, for the CONTENT and PROPERTY relations, the head noun contributes more to NC interpretation. Unexpectedly, for EQUATIVE, the head noun contributes more than the modifier, although only 9 examples were tagged with EQUATIVE, such that the result shown may not be very representative of the general behavior.

8 Discussion

We have presented a method for interpreting the semantic relations of novel NCs using word similarity. We achieved about 53% interpretation accuracy using a path-based measure of similarity. Since our system was tested over raw test data from a general domain, we demonstrated that word similarity has surprising potential for interpreting the semantic relations of NCs. We also investigated using different weights for the head noun and modifier to find out how much the modifier and head noun contributes in NC interpretation and found that, with the exception of some isolated semantic relations, their relative contribution is equal.

Our method has advantages such its relative simplicity and ability to run over small amounts of training data, but there are also a few weaknesses. The main bottleneck is the availability of training data to use in classifying test instances. We suggest that we could use a bootstrap method to overcome this problem: in each step of classification, NCs which are highly similar to training instances, as determined by some threshold on similarity, are added to the training data to use in the next iteration of classification. One way to arrive at such a threshold is to analyze the relative proportion of correctly- and incorrectly-classified instances at different similarity levels, through cross-validation over the training data. We generate such a curve for the test data, as detailed in Figure 6.

If we were to use the crossover point (similarity ≥ 0.57), we would clearly “infect” the training data with a significant number of misclassified instances, namely 30.69% of the new training instances; this would have an unpredictable impact on classification performance. On the other hand, if we were to select a higher threshold based on a higher estimated proportion of correctly-classified instances (e.g. 70%), the relative

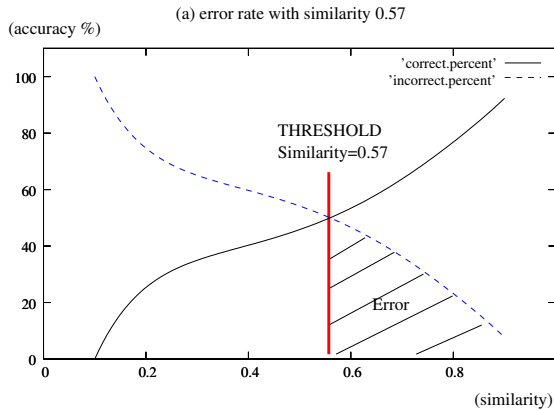


Fig. 6. The relative proportion of correctly- and incorrectly-classified NCs at different similarity values, and the estimated impact of threshold-based bootstrapping

increase in training examples would be slight, and there would be little hope for much impact on the overall classifier accuracy. Clearly, therefore, there is a trade-off here between how much training data we wish to acquire automatically and whether this will impact negatively or positively on classification performance. We leave investigation of this trade-off as an item for future research. Interestingly, in Figure 6 the proportion of misclassified examples is monotonically decreasing, providing evidence for the soundness of the proposed similarity-based model.

In the first experiment (where the weight of the modifier and head noun was the same), we observed that some of the test NCs matched with several training NCs with high similarity. However, since we chose only the NC with the highest similarity, we ignored any insight other closely-matching training NCs may have provided into the semantics of the test NC. One possible workaround here would be to employ a voting strategy, for example, in taking the k most-similar training instances and determining the majority class amongst them. Once again, we leave this as an item for future research.

Acknowledgements

We would like to express our thanks to Bharaneedharan Rathnasabapathy for helping to tag the noun compound semantic relations, and the anonymous reviewers for their comments and suggestions.

References

1. Cao, Y., Li, H.: Base noun phrase translation using web data and the em algorithm. In: COLING2002. (2002)
2. Baldwin, T., Tanaka, T.: Translation by machine of compound nominals: Getting it right. In: ACL2004-MWE, Barcelona, Spain (2004) 24–31

3. Lauer, M.: Designing Statistical Language Learners: Experiments on Noun Compounds. PhD thesis, Macquarie University (1995)
4. Lapata, M.: The disambiguation of nominalizations. *Comput. Linguist.* **28** (2002) 357–388
5. Finin, T.W.: The semantic interpretation of compound nominals. PhD thesis, University of Illinois, Urbana, Illinois, USA (1980)
6. Vanderwende, L.: Algorithm for automatic interpretation of noun sequences. In: Proceedings of the 15th conference on Computational linguistics. (1994) 782–788
7. Rosario, B., Marti, H.: 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. (2001) 82–90
8. Moldovan, D., Badulescu, A., Tatu, M., Antohe, D., Girju, R.: Models for the semantic classification of noun phrases. *HLT-NAACL 2004: Workshop on Computational Lexical Semantics* (2004) 60–67
9. Fan, J., Barker, K., Porter, B.W.: The knowledge required to interpret noun compounds. In: Seventh International Joint Conference on Artificial Intelligence. (2003) 1483–1485
10. Barker, K., Szapkowicz, S.: Semi-automatic recognition of noun modifier relationships. In: Proceedings of the 17th international conference on Computational linguistics. (1998) 96–102
11. Artilles, J., Penas, A., Verdejo, F.: Word sense disambiguation based on term to term similarity in a context space. In: *Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*. (2004) 58–63
12. Patwardhan, S., Banerjee, S., Pedersen, T.: Using measures of semantic relatedness for word sense disambiguation. In: Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics. (2003)
13. Resnik, P.: Disambiguating noun groupings with respect to wordnet senses. In: Proceedings of the 3rd Workshop on Very Large Corpus. (1995) 77–98
14. Wu, Z., Palmer, M.: Verb semantics and lexical selection. In: 32nd. Annual Meeting of the Association for Computational Linguistics. (1994) 133–138
15. Fellbaum, C., ed.: *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, USA (1998)
16. Leacock, C., Chodorow, N.: Combining local context and wordnet similarity for word sense identification. [15]
17. Jiang, J., Conrath, D.: Semantic similarity based on corpus statistics and lexical taxonomy. In: Proceedings on International Conference on Research in Computational Linguistics. (1998) 19–33
18. Lin, D.: An information-theoretic definition of similarity. In: Proceedings of the International Conference on Machine Learning. (1998)
19. Banerjee, S., Pedersen, T.: Extended gloss overlaps as a measure of semantic relatedness. In: Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence. (2003) 805–810
20. Levi, J.: The syntax and semantics of complex nominals. In: New York: Academic Press. (1979)
21. Downing, P.: On the creation and use of English compound nouns. *Language* **53** (1977) 810–42

Appendix

Table 5. The Semantic Relations in Noun Compounds (N_1 = modifier, N_2 = head noun)

<i>Relation</i>	<i>Definition</i>	<i>Example</i>	<i># of test/training</i>
AGENT	N_2 is performed by N_1	<i>student protest, band concert</i>	10(2)/5
BENEFICIARY	N_1 benefits from N_2	<i>student price, charitable compound</i>	10(1)/7(1)
CAUSE	N_1 causes N_2	<i>printer tray, flood water</i>	54(10)/74(11)
CONTAINER	N_1 contains N_2	<i>exam anxiety</i>	13(6)/19(5)
CONTENT	N_1 is contained in N_2	<i>paper tray, eviction notice</i>	40(5)/34(7)
DESTINATION	N_1 is destination of N_2	<i>game bus, exit route</i>	2(1)/2
EQUATIVE	N_1 is also head	<i>composer arranger, player coach</i>	9/17(3)
INSTRUMENT	N_1 is used in N_2	<i>electron microscope, diesel engine</i>	6/11(2)
LOCATED	N_1 is located at N_2	<i>building site, home town</i>	12(2)/16(4)
LOCATION	N_1 is the location of N_2	<i>lab printer, desert storm</i>	29(10)/24(5)
MATERIAL	N_2 is made of N_1	<i>carbon deposit, gingerbread man</i>	12(1)/15(2)
OBJECT	N_1 is acted on by N_2	<i>engine repair, horse doctor</i>	88(16)/88(21)
POSSESSOR	N_1 has N_2	<i>student loan, company car</i>	32(3)/22(4)
PRODUCT	N_1 is a product of N_2	<i>automobile factory, light bulb</i>	27(1)/32(9)
PROPERTY	N_2 is N_1	<i>elephant seal</i>	76(5)/85(7)
PURPOSE	N_2 is meant for N_1	<i>concert hall, soup pot</i>	160(23)/160(23)
RESULT	N_1 is a result of N_2	<i>storm cloud, cold virus</i>	7(4)/8(1)
SOURCE	N_1 is the source of N_2	<i>chest pain, north wind</i>	86(21)/99(18)
TIME	N_1 is the time of N_2	<i>winter semester, morning class</i>	26(2)/19
TOPIC	N_2 is concerned with N_1	<i>computer expert, safety standard</i>	465(51)/446(60)

The 4th column gives us the number of words tagged with the corresponding relation in the 1st column. The numbers within the parenthesis gives us the number of words that are tagged with multiple relations(i.e. those that are tagged with the relation in the 1st column and other relations as well). In the training data, 94 NCs have multiple relations and in test data, 81 NCs have multiple relations.