Assigning Verbs to Semantic Classes via WordNet

Anna Korhonen
University of Cambridge, Computer Laboratory
15 JJ Thomson Avenue, Cambridge CB3 0FD, UK
Anna.Korhonen@cl.cam.ac.uk

Abstract

We propose a method for semi-automatic classification of verbs to Levin classes via the semantic network of WordNet. The method involves first classifying entire WordNet senses to semantic classes and then classifying individual verbs on the basis of their WordNet senses. We report evaluation which shows that the method can be used to build a verb classification accurate enough for practical NLP use. The WordNet-Levin mapping produced as a by-product, may, in turn, be used to supplement WordNet with novel information.

1 Introduction

Linguistic research has shown that verbs fall into classes distinctive in terms of their syntactic and semantic properties (Jackendoff, 1990; Hale and Keyser, 1993; Levin, 1993; Pinker, 1989). For example, verbs which share the meaning component of 'motion' (e.g. *fly* and *walk*) tend to behave similarly also in terms of subcategorization and can thus be grouped to a linguistically coherent class.

While the correspondence between the syntax and semantics of verbs is arguably not perfect and while the whole notion of a verb class is somewhat elusive[1], verb classifications can be constructed which provide a generalization over a range of syntactic and semantic properties of verbs.

Such classifications are particularly useful from a practical NLP point of view. They can be used as a means of reducing redundancy in the lexicon and for filling gaps in lexical knowledge. To a certain extent, they enable inferring the semantics of a word on the basis of its syntactic behaviour, and the syntax of a word on the basis of its semantic behaviour.

Verb classifications have, in fact, been used to support various NLP tasks, including machine translation, language generation (Dorr, 1997), document classification (Klavans and Kan, 1998), lexicography (Sanfilippo, 1994) and lexical acquisition, such as word sense disambiguation (Dorr and Jones, 1996) and subcategorization acquisition (Korhonen, 2002b).

The verb classification employed most widely in NLP is Levin's taxonomy of verbs and their classes (Levin, 1993). Levin classes are based on the ability of a verb to occur in specific diathesis alternations, i.e. specific pairs of syntactic frames which are assumed to be meaning retentive. The classification covers a substantial number of diathesis alternations occurring in English. It is not, however, exhaustive. More work is required on extending and refining it until a comprehensive resource can be obtained suitable for large-scale NLP use (Dorr and Jones, 1996; Dorr, 1997; Korhonen, 2002b).

Perhaps the most challenging task is to extend the classification with new participants. Manual classification of verbs to semantic classes yields accurate results but is time consuming (Levin, 1993; Dang et al., 1998). Fully automatic classification, on the other hand, is fast but suffers from low accuracy (Dorr, 1997; Stevenson and Merlo, 1999). In this paper, we propose combining the strengths of these approaches. We present a method for semi-automatic semantic classification of verbs which exploits automatic techniques but also allows for some manual intervention.

The method involves classifying verbs semantically via the semantic network of WordNet (Miller, 1990). Unlike Levin's source, Wordnet

[1] For example, as most verbs can be characterized by several meaning components, there is potential for cross-classification. Therefore different, equally viable classification schemes can be constructed.

is a comprehensive lexical database. Although it classifies verbs on a purely semantic basis, the syntactic regularities studied by Levin are to some extent reflected by semantic relatedness as it is represented by WordNet's particular structure (Dorr, 1997; Fellbaum, 1999). Our method makes use of this partial overlap between WordNet and Levin classes. It involves first classifying entire WordNet senses to semantic classes and then classifying individual verbs on the basis of their WordNet senses.

We use this method to classify verbs in a number of WordNet files and report experimental evaluation which shows that the method is fairly accurate and can also be used to build a classification suitable for practical NLP use. The classification built using our method can also be used to benefit linguistic research, as knowledge about novel verb and verb class associations can be used to test and enrich linguistic theory (e.g. (Levin, 1993)).

As a by-product, the method produces a mapping between WordNet senses and Levin classes. This mapping – once comprehensive – can act as a useful supplement to WordNet which incorporates information about diathesis alternations and semantic verb classes. Currently this information is absent in WordNet.

We discuss the background for our work in section 2. In section 3, we describe the method for classifying verbs semantically via WordNet. The details of the experimental evaluation of our method are supplied in section 4. Section 5 concludes with directions for future work.

2 Background

In Levin (1993), verbs which display the same or similar set of diathesis alternations in the realization of their argument structure are assumed to share certain meaning components and are organized into a semantically coherent class. For instance, the Levin class of "*Break* Verbs" (class 45.1) refers to actions that bring about a change in the material integrity of some entity. It is characterized by its participation (1-3) or non-participation (4-6) in the following alternations:

1. **Causative/inchoative alternation**:
   Tony broke the window ↔ *The window broke*
2. **Middle alternation**:
   Tony broke the window ↔ *The window broke easily*
3. **Instrument subject alternation**:
   Tony broke the window with the hammer ↔ *The hammer broke the window*
4. * ***With/against* alternation**:
   Tony broke the cup against the wall ↔ **Tony broke the wall with the cup*
5. ***Conative alternation**:
   Tony broke the window ↔ **Tony broke at the window*
6. ***Body-Part possessor ascension alternation**:
   **Tony broke herself on the arm* ↔ *Tony broke her arm*

Levin's classification is attractive in providing a summary of the variety of theoretical research done on semantic verb classes and a reference work extensive enough for NLP research. It is not, however, comprehensive in breadth or depth of coverage. It provides a classification of around 3200 verbs into 48 classes (some of which are split further into more distinctive subclasses, making the total number of 191 classes) according to their participation in 79 alternations involving NP and PP complements only.

Work on refining the existent semantic classes and composing novel ones is under way. Dang et al. (1998), for example, have refined the current classification by creating intersective classes for those verbs which share membership of more than one Levin class. Dorr (1997) has created new classes for verbs whose syntactic behaviour differs from the syntactic description of existing Levin classes. Korhonen (2002b) has proposed novel alternations not covered by Levin – particularly those involving sentential complements – with the aim to create new classes for further verb types.

Attempts have also been made to extend the classification with new participants. Levin, for instance, associates only 13 verbs with the class of "*Break* Verbs" (*break, chip, crack, crash, crush, fracture, rip, shatter, smash, snap, splinter, split,* and *tear*), while it is possible that other English verbs fall into this class as well.

Stevenson and Merlo (1999) and Dorr (1997) have proposed methods for automatically infer-

ring participation in verb classes. Stevenson and Merlo have applied machine learning techniques to automatically classify a set of verbs, based on distributional features extracted from a corpus. They achieve 70% accuracy on the task, but their method is applicable to three verb classes only. Dorr has proposed a method applicable to all Levin classes which relies heavily on lexical resources. It uses grammatical information in the LDOCE dictionary (Procter, 1978) alongside semantic information in WordNet. The method yields 61% accuracy in assigning verbs to (one of their) correct classes.

Although fully automatic classification would be an ideal solution, the methods proposed so far have not led to a resource accurate enough for all-purpose practical NLP use (e.g. (Korhonen, 2002b)). On the other hand, manual classification is not an ideal solution as it is time consuming and thus costly. Here, we address this problem by proposing a semi-automatic method which uses automatic techniques to speed up the classification task but allows for some manual intervention to maximise accuracy of classification.

3 Classifying Verbs Semantically via WordNet

Our method involves assigning verbs to semantic classes via WordNet. Section 3.1 provides a brief introduction to this semantic network and discusses its suitability for the task. Section 3.2 gives details of the classification algorithm.

3.1 WordNet

In contrast to Levin's classification, WordNet organizes words on a purely semantic basis without regard to their syntactic properties. Its organization is that of a network of interlinked nodes representing word meanings. The nodes are sets of synonym sets ('synsets'), which consist of all the word forms that can express a given concept (e.g. {*snooze, drowse, doze*}). While synonymy links individual words within synsets, the super-/subordinate relation links entire synsets, resulting in hierarchical structures.

WordNet has several qualities which make it suitable for our purposes. Unlike Levin's source, it is a comprehensive lexical database. It also includes useful information about semantic relations and the frequency of word senses which can be exploited in the classification task. Although WordNet's semantic organization does not always go hand in hand with syntactic information, Dorr and Jones (1996) and Dorr (1997) have demonstrated that synonymous verbs in WordNet exhibit syntactic behaviour similar to that characterised in the classification system of Levin. This enables association of verbs with semantic classes on the basis of their WordNet synonyms.

We utilized the verb hierarchy of WordNet version 1.6. which contains 10,319 word forms whose 22,066 senses spread over 12,127 synsets. Most synsets in this hierarchy are interrelated by a pointer standing for a manner relation troponymy[2]. They divide into 15 subhierarchies which represent different, mostly semantically-driven domains (e.g. "verbs of motion", "verbs of perception").

3.2 Classification Algorithm

In addition to WordNet, our classification algorithm employs the following set of lexical resources:

- Levin's semantic classification of 3200 verbs
- The LDOCE dictionary
- Dorr's (1997) source of LDOCE grammatical codes for Levin classes. This source was obtained by extracting automatically basic syntactic patterns from all the sentences in Levin's book, mapping these onto LDOCE grammatical codes and grouping them into canonical and prohibited codes for each class.

Dorr (1997) employed the same set of lexical resources in her fully automatic classification approach. In Dorr's approach, each unknown verb was assigned to one of its semantic classes by examining the verb's synonyms from WordNet and selecting those whose Levin class is associated with syntactic information matching that of the unknown verb. The notion of

[2]For example, the synset {*snooze, drowse, doze*} is represented as one of the troponyms (subordinates) of the hypernym (superordinate) synset {*rest, repose*}, since *snooze, drowse* or *doze* mean to *rest* or *repose* in a particular manner.

"match" was based on the degree of intersection between the verb's LDOCE codes and the LDOCE codes for a candidate class.

Our semi-automatic approach differs, however, essentially from Dorr's. In our approach, entire WordNet senses (i.e. synsets) are first assigned to semantic classes, using information in the lexical resources merely as a guidance. After synsets are classified, individual verbs receive the classification of their respective synsets. By classifying entire synsets, our objective is to build a more static source where WordNet senses are associated with different Levin classes. Although static, the source will allow for updating and adding new verbs to WordNet.

Our algorithm makes use of the hierarchical structure of WordNet. It classifies synsets subhierarchy by subhierarchy, starting from the top level synsets, and going further down in the taxonomy only when required. This makes sense, as many of the top level synsets intersect directly with Levin classes. For example, "Sending and Carrying" and "Force Exerting" verbs are all found under the same top level synset {*move, displace*}.

Each synset is classified by first assigning the majority of its member verbs to a semantic class and then choosing the Levin class supported by the highest number of verbs. 'Member verbs' refer here to those which are members of the synset in question and of its hyponym (subordinate) synsets. Thus if a classified synset has hyponym synsets, the latter are classified according to their classified hypernym synset.

At the present state of development, our algorithm classifies verbs according to their predominant (i.e. most frequent) sense in WordNet only[3]. Accordingly, synsets are classified by considering only those member verbs whose predominant sense belongs to the synset in question. The method is, however, generally applicable (i.e. fully suitable for classifying non-predominant senses of verbs as well).

The algorithm proceeds as follows:

Step 1: If the majority of member verbs of a given synset S are Levin verbs from the same class, classify S directly.

Step 2: Otherwise, classify more member verbs (according to Step 4a-d) until the majority are classified, and then go back to Step 1.

Step 3: Otherwise, if the classified verbs point to different Levin classes, examine whether S consists of hyponym synsets. If not, assign S to the Levin class supported by the highest number of classified verbs. If yes, go one level down in the hierarchy and classify the hyponym synsets separately, starting again from Step 1.

Step 4: If S includes no Levin verbs, proceed as follows to classify the majority of member verbs of S:

- **(a)** Extract the predominant sense of a given verb V from WordNet
- **(b)** Extract the syntactic codes from LDOCE relevant to this sense
- **(c)** Examine whether V could be assigned to a Levin class already associated with the other verbs in the (i) same synset, (i) possible hypernym synset or (iii) possible sister synsets by comparing the LDOCE codes of the sense and Dorr's LDOCE codes of the respective Levin class(es). Given the hypothesised classes, make the final class assignment manually.
- **(d)** If no suitable class is found, re-examine the case after more verbs have been analysed. If the classification remains unsolved, set V aside for later examination.

The above algorithm is for the most part automatic, however, Step 4b and part of Step 4c (the final class assignment) are done manually to ensure accuracy of classification. The approach of classifying "only" the majority of verbs in synsets (as opposed to all verbs) aims, however, to minimise manual effort.

The following examples illustrate the use of this algorithm to assign hyponym synsets of the top level synset {*move, displace*} to Levin classes[4]:

[3]The WordNet frequency data is derived from the SemCor corpus.

[4]As we only consider verbs whose predominant senses belong to the synsets in question, for clarity, we refer to synsets in these examples as WordNet synset identifier

Example 1: The synset no. 01328437 has 5 member verbs, 3 of which are Levin verbs from the same verb class. It is assigned directly to the Levin class of "Verbs of Sending and Carrying" (Step 1).

```
ship => Verbs of Sending and Carrying
despatch
dispatch => Verbs of Sending and Carrying
route
forward => Verbs of Sending and Carrying
```

Example 2: The synset no. 01278717 includes Levin verbs which point to different classes. Since it consists of hyponym synsets (as indicated by the synset identifiers below), we go one level down in the taxonomy and classify the hyponym synsets separately (Step 3).

```
push => Verbs of Exerting Force
jab poke 01296169 => Poke Verbs
nudge prod 00838894 => Verbs of Contact
repel 01034588
shove 01278320 => Verbs of Exerting Force
ram 01296169
obtrude 01279473
thrust 01296169 => Verbs of Exerting Force
elbow shoulder 01278320
```

Example 3: The synset no. 00994853 includes 13 member verbs, 4 of which are Levin "Verbs of Sending and Carrying". We need to classify more verbs to determine class assignment (Step 2). We choose *whisk* and extract its predominant sense from WordNet (Step 4a):

```
whisk -- (move somewhere quickly;
          "The president was whisked away
          in his limo")
```

In LDOCE the verb has three senses. That corresponding to the predominant WordNet sense is identified as (Step 4b):

```
2. [X9 esp. OFF, AWAY] to remove
   b. by taking suddenly:
      "She whisked the cups away
      / whisked him (off) home"
```

11 Levin classes are already matched with the verbs in the same, hypernym and sister synsets. Those whose syntactic description includes the LDOCE code X9 are:

```
Verbs of putting
Verbs of removing
Verbs of sending and carrying
Verbs of exerting force
Verbs of motion
```

After verifying these options manually, *whisk* is assigned to "Verbs of Sending and Carrying" (Step 4c).

Example 4: The synset no. 01527059 includes around 90 member verbs related to the transfer of messages. These spread over nearly 60 hyponym synsets. Seven of the verbs are Levin verbs from various classes which include verbs taking sentential complements. Two of them are listed by Dorr (1997) as members of her new semantic classes. The synset is set aside for future work (Step 4d).

4 Experimental Evaluation

We applied the verb classification algorithm to three WordNet subhierarchies: contact, possession and motion verbs. 1581 synsets in these files were assigned to one of 32 Levin classes and 148 were left unclassified. A small number of synsets (35) from other verb files were classified as well.

To evaluate the classification algorithm, we chose 60 synsets from among the 1616 classified. The synsets were chosen at random so that 20 were taken from each of the three WordNet verb files (contact, possession and motion verbs). From the total of 501 verbs in these synsets, we chose for evaluation those 224 which were neither Levin verbs nor classified manually when linking synsets with Levin classes. For these verbs, the semantic classification (to a broad or – where applicable – to a fine-grained

codes, rather than their actual names. To simplify the examples somewhat, we also refer to broad rather than fine-grained Levin classes.

Levin class) was compared against a manually obtained gold standard classification.

The algorithm classified correctly 181 verbs and incorrectly 43 verbs. Accuracy of the class assignment was thus 81%.[5] All the misclassifications corresponded to cases where there is a semantic mismatch between WordNet and Levin, i.e. the majority, but not all verbs in a synset corresponded to the same Levin class. 21% of the misclassifications concerned synsets which divide further into hyponym synsets. It is possible that better results could be obtained for these synsets if the classification algorithm was refined further (e.g. to classify more than "just" the majority of the member verbs in a synset, for example 70%, before classifying the entire synset).

79% of the misclassifications concerned, however, synsets which do not divide further into hyponym synsets and whose analysis cannot be improved, if we maintain the idea of linking entire synsets with Levin classes. In the light of these results, it thus seems that the upper bound of our method is around 85%.

The correspondence between WordNet synsets and Levin classes varies, however, considerably across WordNet and this affects the accuracy of our method. Further evaluation is therefore needed with a larger set of WordNet files to get a better idea of the performance of the algorithm.

Our results are not directly comparable with those of Dorr (1997), as Dorr does not restrict her evaluation to a certain sense. She reports 61% accuracy in assigning a verb to *any* of its correct classes, while we report 81% accuracy in assigning a verb to the class corresponding to its predominant sense.

Task-based evaluation of our classification algorithm was reported in (Korhonen, 2002a). In this evaluation, a classification obtained using our algorithm was employed in the task of automatic subcategorization acquisition. The algorithm was adapted to classify verbs to (mainly) broad Levin classes as in subcategorization acquisition it was adequate to assume a broad notion of a semantic class.

The method for subcategorization acquisition involved first using the system of Briscoe and Carroll (1997) to acquire a putative subcategorization frame (SCF) distribution for a verb from corpus data. The system used a shallow parser to obtain the subcategorization information and employed a classification of 163 verbal SCFs. The putative distribution was then smoothed with the back-off distribution corresponding to the semantic class of (the predominant sense of) a verb. The semantic class was determined by our classification algorithm.

This task-based evaluation is fairly sensitive to the accuracy of the class assigments. Where the predominant sense is assigned correctly, significant improvement can be reported in SCF acquisition, as accurate back-off estimates help to correct the acquired SCF distribution and deal with sparse data. Incorrect class assignments can, however, degrade performance.

On a test set of 91 verbs from 16 Levin classes, the method yielded 81% type precision (the percentage of SCF types that the method proposes which are correct) and 73% type recall (the percentage of SCF types in the gold standard that the method proposes). The method was compared against a baseline method which assumed no semantic class and involved no smoothing. F measure[6] was 76 for the method that assumed a semantic class and 61 for the baseline. This result shows that our algorithm is accurate enough to yield a classification which can benefit a practical NLP task.

5 Conclusion

We have proposed a method for semi-automatic semantic classification of verbs which combines the benefits of manual classification and automatic techniques. The method assigns verbs to semantic classes via WordNet, by first classifying entire WordNet senses and then classifying individual verbs according to their WordNet senses. We have reported evaluation which shows that method is fairly accurate and can be used to build a classification suitable for practical NLP use.

In the future, we will apply the method to other Wordnet files, perform more comprehensive evaluation and further refine the

[5] Korhonen (2002a) reported a smaller evaluation of this classification algorithm with 151 verbs which yielded 93% accuracy. However, in this evaluation a broad notion of Levin class was assumed which explains the higher accuracy.

[6] $F = \frac{2 \cdot precision \cdot recall}{precision + recall}$

classification algorithm to consider also non-predominant senses of verbs. Our long term goal is to build a comprehensive semantic classification of verbs which can be useful for NLP as well as for linguistic research and – as a by-product – a mapping between WordNet senses and Levin classes which can act as a useful supplement to WordNet.

6 Acknowledgements

We thank Bonnie Dorr for the use of her source of LDOCE grammatical codes for Levin classes. This work was partly supported by UK EPSRC project GR/N36462/93: 'Robust Accurate Statistical Parsing (RASP)'.

References

E. J. Briscoe and J. Carroll. 1997. Automatic extraction of subcategorization from corpora. In *5th ACL Conference on Applied Natural Language Processing*, pages 356–363, Washington DC.

H. T. Dang, K. Kipper, M. Palmer, and J. Rosenzweig. 1998. Investigating regular sense extensions based on intersective Levin classes. In *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, pages 293–299, Montreal, Canada.

B. Dorr and D. Jones. 1996. Role of word sense disambiguation in lexical acquisition: predicting semantics from syntactic cues. In *16th International Conference on Computational Linguistics*, pages 322–333, Copenhagen, Denmark.

B. Dorr. 1997. Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Translation*, 12(4):271–325.

C. Fellbaum. 1999. The organization of verbs and verb concepts in a semantic net. In P. Saint-Dizier, editor, *Predicative Forms in Natural Language and in Lexical Knowledge Bases*, pages 93–110. Kluwer Academic Publishers, Netherlands.

K. Hale and S. J. Keyser. 1993. *On Argument Structure and Lexical Expression of Syntactic Relations*. MIT Press, Cambridge, Massachusetts.

R. Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge, Massachusetts.

J. L. Klavans and M. Kan. 1998. Role of verbs in document analysis. In *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, pages 680–686, Montreal, Canada.

A. Korhonen. 2002a. Semantically motivated subcategorization acquisition. In *ACL Workshop on Unsupervised Lexical Acquisition*, Philadelphia. To appear.

A. Korhonen. 2002b. *Subcategorization Acquisition*. Ph.D. thesis, University of Cambridge, UK.

B. Levin. 1993. *English Verb Classes and Alternations*. Chicago University Press, Chicago.

G. A. Miller. 1990. WordNet: An on-line lexical database. *International Journal of Lexicography*, 3(4):235–312.

S. Pinker. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. MIT Press, Cambridge, Massachusetts.

P. Procter. 1978. *Longman Dictionary of Contemporary English*. Longman, England.

A. Sanfilippo. 1994. Word knowledge acquisition, lexicon construction and dictionary compilation. In *International Conference on Computational Linguistics*, pages 273–277, Kyoto, Japan.

S. Siegel and N. J. Castellan, editors. 1988. *Non-Parametric Statistics for the Behavioural Sciences*. McGraw-Hill, New York.

S. Stevenson and P. Merlo. 1999. Automatic verb classification using distributions of grammatical features. In *9th Conference of the European Chapter of the Association for Computational Linguistics*, pages 45–52, Bergen, Norway.