Combining Contexts in Lexicon Learning for Semantic Parsing

Richard Socher
Saarland University
richard@socher.org

Chris Biemann
University of Leipzig
Ifi, NLP Department
Johannisgasse 26
04103 Leipzig, Germany
biem@informatik.uni-leipzig.de

Rainer Osswald
FernUniversität in Hagen
Department of Mathematics
and Computer Science, IICS
rainer.osswald@fernuni-hagen.de

Abstract

We introduce a method for the automatic construction of noun entries in a semantic lexicon. Using the entries already present in the lexicon, semantic features are inherited from known to yet unknown words along similar contexts. As contexts, we use three specific syntactic-semantic relations: modifying adjective, verb-deep-subject and verb-deep-object. The combination of evidences from different contexts yields very high precision for most semantic features, giving rise to the fully automatic incorporation into the lexicon.

1 Introduction

Advanced tasks such as text summarization and question answering call for tools that support the semantic analysis of natural language texts. While syntactic parsers have been intensively studied for decades, broad coverage semantic parsing is a relatively recent research topic. Semantic parsing aims at constructing a semantic representation of a sentence, abstracting from the syntactic form and allowing queries for meaning rather than syntax. That is, semantic relations between concepts are in the focus of interest rather than syntactic relations between words and phrases.

A major current line of research for extracting semantic structures from texts is concerned with semantic role labeling. The FrameNet database (Baker et al., 1998) provides an inventory of semantic frames together with a list of lexical units associated with these frames. Semantic parsing then means to choose appropriate semantic frames from the frame inventory depending on the lexical concepts present in the given sentence and to assign frame-specific roles to concepts. A related task has been defined as part of CoNLL 2004 (Carreras and Màrques, 2004). Here, machine learning methods are used to learn a semantic role labeler from an annotated text to extract a fixed set of semantic relations.

If one aims at deep semantic parsing, a lexicon containing semantic information about words and concepts is a prerequisite. Building such a lexicon is a time-consuming and expensive task. The acquisition bottleneck is extremely thin in this area, as lexicon entries tend to be rather complex. Therefore, methods that are capable of automatically or semi-automatically extending semantic lexicons are highly needed to overcome the bottleneck and to scale the lexicon to a size where satisfactory coverage can be reached. In this paper, we present a method that enlarges the number of noun entries in the lexicon of a semantic parser for German.

1.1 Related Work

Extending a given lexicon with the help of a parser relying on this lexicon can be viewed as a step of a bootstrapping cycle: Lexicon entries of known words are used to obtain entries for previously unknown words by exploiting a parsed corpus.

Early bootstrapping approaches such as (Riloff and Shepherd, 1997) were based on few seed words of a semantic category and their nearest neighbor contexts. Higher precision was achieved by separating extraction patterns into two groups by (Roark

Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek and Mare Koit (Eds.)
NODALIDA 2007 Conference Proceedings, pp. 175–182

and Charniak, 1998): Conjunctions, lists and appositives being one and noun compounds being the other. Since bootstrapping single categories often leads to category shifts in later steps, (Thelen and Riloff, 2002) use an un-annotated corpus, seed words and a large body of extraction patterns to discover multiple semantic categories like *event* or *human* simultaneously.

Another related research line is distributional clustering to obtain semantic classes via similar contexts, e.g. (Pereira et al., 1993; Lin, 1998; Rooth et al., 1999). Here, semantic classes are created by the clustering method rather than assigned to predefined classes in the lexicon; these works also employ the distributional hypothesis, i.e. that similar semantic properties are reflected in similar (syntactic) contexts.

Our setup differs from these approaches in that we use a conceptual framework that covers all sorts of nouns rather than concentrating on a small set of domain-specific classes or leaving the definition of the classes to the method. Moreover, we combine three different context types; see (Hagiwara et al., 2006) for a discussion on context combination for synonym acquisition.

2 Semantic Lexicon and Parser

This section gives a brief outline of the semantic parsing framework with respect to which our learning task is set up. The task is to automatically extend a semantic lexicon used for semantic parsing by exploiting parses that have been generated on the basis of the already existing lexicon. From these parses we extract three types of syntactic-semantic noun contexts which are then employed to classify unknown nouns on the basis of classified nouns, as explained in more detail in Section 3.

2.1 The MultiNet Formalism

The semantic parses we exploited for our experiments comply with the MultiNet knowledge representation formalism (Helbig, 2006). MultiNet represents the semantics of natural language expressions by means of semantic networks, where nodes represent concepts and edges represent relations between concepts. Each concept node is labeled by an element from a predefined upper-domain hierarchy of

object [*o*]				
	concrete object [*co*]			
		discrete object [*d*]		*house, apple, tiger*
		substance [*s*]		*milk, honey, iron*
	abstract object [*ab*]			
		attribute [*at*]		
			measurable attribute [*oa*]	*weight, length*
			non-measurable attribute [*na*]	*form, trait, charm*
		relationship [*re*]		*causality, similarity*
		ideal object [*io*]		*justice, category*
		abstract temporal object [*ta*]		*Easter, holiday*
		modality [*mo*]		*necessity, permission*
		situational object [*abs*]		
			dynamic situational object [*ad*]	*race, robbery*
			static situational object [*as*]	*equilibrium, sleep*
quantity [*qn*]				
	unit of measurement [*me*]			*kg, meter, mile*
	...			

Table 1: Part of the MultiNet sort hierarchy relevant to concepts expressed by nouns

45 *ontological sorts* such as ‘discrete object’ (*d*), ‘attribute’ (*at*), and ‘situational object’ (*abs*) (see Table 1) and various so-called layer features indicating facticity, quantification, and referential determination among other things. In addition, MultiNet comprises about one hundred *semantic relations* including a set of *semantic case roles* such as AGT (agent), AFF (affected object), MEXP (mental experiencer) as well as relations for expressing causation, implication, temporality, and so on. The reader is referred to (Helbig, 2006) for a detailed account of the MultiNet paradigm.

2.2 Semantic Parser

The parsed German corpora used in our experiments have been produced by the syntactic-semantic parser described in (Hartrumpf, 2003). This parser, which has been successfully employed for information retrieval and question answering tasks (Hartrumpf, 2005), relies on the computational lexicon HaGenLex (see below) and has components for word sense disambiguation, compound analysis, and coreference resolution. In addition to MultiNet structures, the parser also generates syntactic dependency trees.

The semantic structures produced by the parser depend essentially on the semantic roles specified in the valency frames of the HaGenLex entries. Working with these semantic parses enables us to investigate specific syntactic-semantic contexts of nouns with respect to their potential to act as indicators for

Name	Meaning	Examples	
		+	−
ANIMAL	animal	*fox*	*person*
ANIMATE	living being	*tree*	*stone*
ARTIF	artifact	*house*	*tree*
AXIAL	object with distinguished axis	*pencil*	*sphere*
GEOGR	geographical object	*the Alps*	*table*
HUMAN	human being	*woman*	*ape*
INFO	(carrier of) information	*book*	*grass*
INSTIT	institution	*UNO*	*apple*
INSTRU	instrument	*hammer*	*lake*
LEGPER	juridical or natural person	*firm*	*animal*
MENTAL	mental object or situation	*pleasure*	*length*
METHOD	method	*procedure*	*book*
MOVABLE	object being movable	*car*	*forest*
POTAG	potential agent	*motor*	*poster*
SPATIAL	object with spatial extension	*table*	*idea*
THCONC	theoretical concept	*category*	*fear*

Table 2: Set of 16 binary semantic features

art-substance

$$\begin{bmatrix} \text{SORT} & s \\ \text{ANIMATE} & - \\ \text{ARTIF} & + \\ \text{INFO} & - \\ \text{MOVABLE} & + \\ \ldots & \end{bmatrix}$$

con-info

$$\begin{bmatrix} \text{SORT} & d \\ \text{ANIMATE} & - \\ \text{ARTIF} & + \\ \text{INFO} & + \\ \text{MOVABLE} & + \\ \ldots & \end{bmatrix}$$

Figure 1: Two examples of complex semantic sorts: *art-substance* (artifical substance) and *con-info* (concrete information object).

the semantic sort of the nouns. In the experiments described in the following, we focus on the *argument position* a noun takes in the valency frame of a verb, thereby abstracting from the specific semantic role of the argument. Though attractive in principle, preliminary investigations have indicated a sparse-data problem for specific semantic roles.

2.3 The Lexicon HaGenLex

In our experiments, we use part of the computational lexicon HaGenLex (Hagen German Lexicon) as training data. HaGenLex contains about 25,000 German lexical entries (13,000 nouns, 7,000 verbs) with detailed morphological, syntactic, and semantic specifications (Hartrumpf et al., 2003). The semantic specification of HaGenLex entries rests on the MultiNet formalism, that is, every entry is assigned with an ontological sort of the MultiNet sort hierarchy and all valency frames are equipped with MultiNet case roles.

In addition, the noun entries in HaGenLex are classified with respect to 16 *binary semantic features* such as ANIMATE, HUMAN, ARTIF(ICIAL), and INFO(RMATION); see Table 2 for the full list. The features and ontological sorts are not independent of each other; e.g., HUMAN:+ implies ANIMATE:+, ARTIF:−, and sort *d* (discrete object). To prevent inconsistencies, the possible combinations of semantic features and ontological sorts are explicitly combined into *complex semantic sorts*. Figure 1 shows two examples of such combined sorts: *art-substance* (artifical substance), e.g., *paper*, *beer*, and *con-info* (concrete information object), e.g., *poster*, *certificate*. In total, there are 50 complex semantic sorts; Table 7 lists the 15 most frequent of them in our training data. Since not all of the complex semantic sorts are specified with respect to every feature and because of the ontological sort hierarchy, there is a natural *specialization hierarchy* on the set of complex sorts.

3 Method

The goal of our experiments is to assign complex semantic sorts to unknown nouns. To this end, we separately train binary classifiers for the ontological sorts and semantic features and combine their results in a second step to a complex semantic sort, if possible. Since we use 16 features (Table 2) and 17 sorts (Table 1), this leads to 33 binary classifiers. It should be mentioned that certain classifiers show a fairly strong bias with respect to their distribution within the noun entries of HaGenLex, which in turn gives rise to some unwelcome effects for the respective training results. The *bias* is here defined as the proportion of the more frequent of the two classes.

3.1 Data and Data Structure

Following the distributional hypothesis (Harris, 1968), nouns in a similar context can be assumed to share semantic properties. In our experimental setup, the context of each noun consists of one co-occurring word in a specific relation. These *context elements* are adjectives or verbs in their respective base form. They are automatically disambiguated by the parser if multiple polysemous meanings are present in the lexicon. (The different meanings are indicated by numbers attached to the base form; cf. Figure 2.) In the case of verbs we further distinguish between the different argument positions taken by

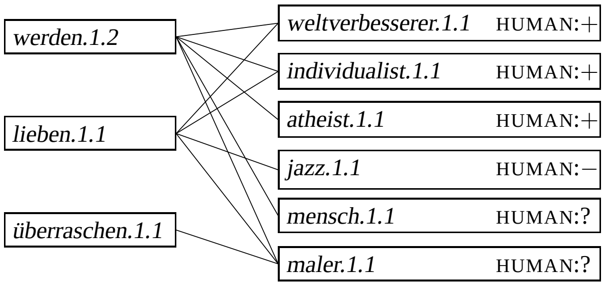

Figure 2: Small sample data for German verb-deep-object relations. Connections indicate co-occurrence in the data.

the noun in the valency frame of the verb. For simplicity, we restrict ourselves to the first two argument positions, henceforth referred to the *deep-subject* and the *deep-object* position, respectively. Notice that according to this terminology, the grammatical subject in passive voice coincides with the deep object.

Pairs of context elements and nouns are aligned in a bipartite graph where known nouns are connected to their context elements which in turn are connected to unknown words, as shown in Figure 2. Before the algorithm starts, a profile of context elements for each noun is extracted from a corpus, which was parsed with the semantic parser described in Section 2.2. It states how often a noun co-occurs with its context elements. We used a corpus of 3,068,945 sentences obtained from the Wortschatz project (Biemann et al., 2004), consisting mainly of contemporary newspaper texts. The parser has a coverage of 42% on this corpus.

3.2 Algorithm

After initialization, the algorithm runs with one specific context type (adjective, deep-subject, or deep-object) at a time. For each context type the process is carried out for all 33 classifiers. (The combination of all context types into one bipartite graph produced worse results in preliminary experiments.) The algorithm's core loop alternates between two phases:

1. *Training*: Each context element gets assigned probabilities that express how indicative this word is for both the positive and the negative class of the classifier in this run. The probability is calculated by dividing the frequency distribution of each class by the total number of nouns in that class, followed by a normalization step per context element.
2. *Classification*: New classes are assigned to unclassified nouns by multiplying the normalized class probabilities of all context elements class-wise from their profiles of context elements. Class probabilities are multiplied only from context elements that had occurred in a profile of a known noun. New nouns get the class with the highest resulting probability.

This alternation of profile calculation and classification is iterated in a bootstrapping fashion. A difference to other bootstrapping methods like those mentioned in the introduction is that the algorithm only iterates about five times, classifying about 95% of all new nouns during the first iteration. The class profiles are updated based on the new classifications and the cycle starts again unless no new nouns are classified. Figure 3 shows the algorithm in pseudo code.

```
Initialize the training set;
While (new nouns get classified){
   Calculate context element profiles;
   For (each unclassified noun n){
      Multiply class probabilities
         class-wise;
   Assign class with highest
      probability to noun n;
   }
}
```

Figure 3: Algorithm for noun classification in pseudo code

Modification of the algorithm is possible by introducing a threshold α for the minimum number of context elements a noun has to co-occur with in order to be assigned their class. Several experiments proved $\alpha = 5$ to be a good heuristics. With fewer evidence precision drops significantly and with higher numbers recall drops without a gain in precision. In the next section, we will illustrate the algorithm with an example classification.

3.3 Example

For a small demonstration we use the data depicted in Figure 2. First the distribution of classes per context element is calculated as shown in Figure 4a.

a)

Context element	HUMAN:+	HUMAN:−
werden.1.2	3	0
lieben.1.1	2	1

b)

werden.1.2	3/3	0/1
lieben.1.1	2/3	1/1

c)

werden.1.2	1	0
lieben.1.1	0.4	0.6

Figure 4: Stepwise calculation of class probabilities per context element for classifier HUMAN.

Then the distribution is divided by the total number of nouns in that class (Figure 4b). Finally, the relative frequencies are normalized to one per context element; e.g., for *lieben.1.1*:

$$P(\text{HUMAN:}+) = \frac{2/3}{2/3+1/1} = 2/5 = 0.4$$

Figure 4c shows the resulting probability vectors. Now the other nouns get classified by combining the probabilities of co-occurring context elements: For *mensch.1.1* the probabilities are: $P(\text{HUMAN:}+) = 1$ and $P(\text{HUMAN:}-) = 0$, so *mensch.1.1* is HUMAN:+ with high confidence. The case for *maler.1.1* is a bit more difficult because it co-occurs with two different context elements, whose probabilities are multiplied class-wise: $P(\text{HUMAN:}+) = 1 \cdot 0.4 = 0.4$ and $P(\text{HUMAN:}-) = 0 \cdot 0.6 = 0$. So *maler.1.1* also gets the class HUMAN:+ because *werden.1.2* does not occur with a HUMAN:− noun. Notice here that the verb *berraschen.1.1* could not be used since it has not yet appeared in any profile of a known noun. This changes in the next iteration, as now *berraschen.1.1* appears in the profile of the newly classified *maler.1.1*. Further notice that *werden.1.2* with probability 0 for HUMAN:− prevents *maler.1.1* to ever get this characteristic. Smoothing, i.e. assigning small minimum probabilities for all classes did, however, not affect the results much in previous experiments and was therefore not undertaken.

3.4 Building Complex Semantic Sorts

After bootstrapping the 33 binary classifiers, their outcomes can be used to build complex semantic sorts. Previous experiments showed significantly better results for single characteristics than executing the method directly on the 50 complex semantic sorts introduced in Section 2.3. The results of the binary classifiers for a given noun are combined as follows:

(1) Determine all complex semantic sorts whose semantic features and ontological sorts are compatible with the results of all binary classifiers.

(2) From the results of (1) select those sorts that are minimal with respect to the specialization relation defined on the set of complex semantic sorts (see Section 2.3).

(3) If the set determined in (2) contains exactly one element, then assign this semantic sort to the given noun, otherwise refuse a classification.

3.5 Combination of Context Types

While past experiments on extending HaGenLex (Biemann and Osswald, 2006) have solely been conducted with modifying adjectives as context elements of nouns, the present study also investigates verb-deep-object and verb-deep-subject relations. It is thus possible to combine the results of different context types. In our experiments, the combination is carried out two ways: In a *lenient* setting, only those nouns are left in, which were assigned the same class (positive or negative) of the same classifier at least twice during the experiments with adjective, deep-subjects and deep-objects. In a *strict* setting, classifications in all three context types had to agree.

By combining the results from different context types we gain stronger evidence for each characteristic, possibly at the cost of losing recall. Since we aim rather at producing correct entries in the lexicon than a bulk of wrong ones, precision is the primary measure to optimize here.

4 Experiments

After parsing the corpus with the semantic parser, we extracted the following numbers of different co-occurrences for each context type:

430,916 verb-deep-subject
408,699 verb-deep-object
450,184 adjective-noun

Context type	Bias	Prec.	Rec.	max.Rec.
adjective-noun	0.841	0.927	0.122	0.390
verb-d-subject	0.850	0.967	0.111	0.339
verb-d-object	0.837	0.973	0.094	0.292

Table 3: Average bias, precision, recall and maximally possible recall of three context types ($\alpha = 5$).

Context type	Precision	Recall
adjective-noun	0.845	0.319
verb-d-object	0.891	0.292
verb-d-subject	0.878	0.248

Table 4: Results with $\alpha = 1$ as a preparation for context type combination.

For evaluation we used 10-fold-cross-validation on 11,100 HaGenLex nouns, where the partition was ensured to retain class distribution.

4.1 Experiments

Table 3 shows the arithmetic means of all 33 characteristics and their respective bias, precision, recall and maximally possible recall for each context type. Note that the maximal recall is bounded by the relatively small intersection of known nouns in the database and nouns in the particular part of the corpus.

The results clearly demonstrate the superiority of verbal contexts in this approach, improving precision by 5% as compared to adjective modifiers and not supporting the findings of (Hagiwara et al., 2006), where the modifier relation was reported to perform best on a related task. Bootstrapping on pairs of verb-deep-object co-occurrences shows a precision of 97.3% averaged over all characteristics.

Nevertheless it seems promising to combine the classifier results from different context types by the methods described in Section 3.5 to classify new nouns more correctly. In this setting, the parameter α was reduced from 5 to 1, that is, only one context element is sufficient for new nouns to be classified. The average results for these single context bootstrapping runs are listed in Table 4. Precision is much lower in these experiments because of the parameter setting. The main objective is a high recall, since high precision is supposed to be created by the combination of these results as described in Section 3.5. The outcome of this process is displayed in Table 5, yielding mostly higher precision and higher recall values than the results of using only a single context type.

Sem. feature or ontol. sort	Strict comb. Prec.	Strict comb. Rec.	Lenient comb. Prec.	Lenient comb. Rec.
HUMAN	0.997	0.143	0.969	0.273
GEOGR	0.982	0.136	0.908	0.257
SPATIAL	0.988	0.123	0.943	0.263
LEGPER	0.997	0.141	0.970	0.275
INSTIT	0.995	0.172	0.972	0.289
ANIMAL	0.996	0.179	0.985	0.298
POTAG	0.995	0.138	0.969	0.274
MOVABLE	0.967	0.120	0.900	0.249
ANIMATE	0.996	0.141	0.970	0.274
INFO	0.967	0.129	0.874	0.243
THCONC	0.944	0.115	0.849	0.238
METHOD	0.997	0.179	0.983	0.300
AXIAL	0.964	0.123	0.904	0.251
MENTAL	0.986	0.159	0.948	0.278
INSTR	0.981	0.143	0.918	0.263
ARTIF	0.926	0.092	0.817	0.217
d	0.980	0.121	0.927	0.257
na	0.996	0.176	0.977	0.297
abs	0.970	0.117	0.907	0.253
mo	0.998	0.182	0.988	0.304
ta	0.989	0.157	0.959	0.279
co	0.988	0.123	0.943	0.263
ab	0.989	0.124	0.945	0.262
s	0.992	0.165	0.965	0.288
oa	0.996	0.176	0.980	0.295
io	0.927	0.096	0.825	0.224
o	1.000	0.191	0.998	0.316
me	1.000	0.191	0.998	0.316
qn	1.000	0.191	0.998	0.316
ad	0.960	0.120	0.891	0.251
at	0.991	0.164	0.956	0.283
re	1.000	0.197	1.000	0.322
as	0.947	0.125	0.860	0.242
Average	0.982	0.147	0.939	0.273

Table 5: Precision and Recall for the combination of classifications using different context types.

Evaluating the combination to complex semantic sorts, the verb-deep-subject contexts gives the best results, as Table 6 indicates. The lenient combination's recall is almost twice as high, but falls short on precision. Notice that in this case average values are not obtained by the arithmetic mean of all semantic sorts, but by the total number of correctly and falsely identified nouns.

Table 7 shows the cross-validation results for assigning complex semantic sorts for the 15 most frequent sorts in the 11,100 noun sample. Examples

Context type	Recall	Precision
adjective	0.039	0.684
deep-object	0.049	0.758
deep-subject	0.057	0.872
lenient combination	0.113	0.772
strict combination	0.059	0.752

Table 6: Recall and precision of complex semantic sorts, with $\alpha = 5$ for single context types and $\alpha = 1$ for combinations

Complex sort	#	%Rec.	%Prec.
nonment-dyn-abs-situation	2752	12.83	88.92
human-object	2359	20.26	94.65
prot-theor-concept	815	1.23	62.50
animal-object	593	0.84	100.00
ax-mov-art-discrete	568	1.06	60.00
plant-object	445	0.22	25.00
nonment-stat-abs-situation	378	1.32	62.50
nonmov-art-discrete	191	3.14	40.00
nonax-mov-art-discrete	174	0.57	16.67
mov-nonanimate-con-potag	159	1.89	50.00
abs-info	148	4.73	53.85
art-substance	147	1.36	50.00
tem-abstractum	143	1.40	100.00
art-con-geogr	138	1.45	28.57
nat-substance	130	1.54	25.00

Table 7: Complex semantic sorts, number of nouns in the initial set, recall and precision.

for the largest group of non-mental dynamic situations are *wettbewerb.1.1* (‘competition’), *zusammenarbeit.1.1* (‘co-operation’), *apokalypse.1.1* (‘apocalypse’) or *aufklï¿½ung.1.2* (‘elucidation’). Only four semantic sorts have a precision above 75%. Recall is only satisfactory for two sorts. In total, 1,041 new nouns (i.e. not listed in the lexicon before) were classified. Various sorts cannot successfully be identified with this method or the used settings. The following list shows all semantic sorts that have not been assigned to any nouns, even though they occurred more than 100 times in the initial set of 11,100 nouns: *nonoper-attribute*, *ment-stat-abs-situation*, *nat-discrete* and *prot-discrete*.

While the number of new nouns for which a combination to complex semantic sorts was possible is not very satisfying, there seems to be room for improvement by exploiting the binary characteristics in a more sophisticated way than by the straightforward algorithm described in Section 3.5. The combined run on the three context types, on which this combination to semantic sorts is based, created 125,491 new single binary characteristics for 3,755 nouns not in the lexicon. It should be possible to improve this number by using a larger corpus.

4.2 Discussion of Results

The presented experiments use fine grained binary features rather than complex semantic sorts as in most other works. Evidence from different relations improves results to an average of 98.2% precision for binary characteristics, with most characteristics above 99% (see Table 5). If the time consuming work of creating a reasonably small lexicon of nouns and their binary characteristics is done once, our algorithm can then be used effectively to increase the lexicon size.

However, some classes with a skewed distribution (bias above 0.8) have the problem of not assigning the smaller class correctly. Almost half of the 33 characteristics have this problem in bootstrapping runs over single context types. With the strict and lenient combination the problem is alleviated since the few wrong classifications rarely occur twice.

In an environment where incomplete semantic specifications are allowed in the lexicon some characteristics can be incorporated without supervision. If precision is the only concern and a small recall is acceptable then only nouns with three identical values (98.2% precision, 14.7% recall) should be used, whereas results with two identical values return more new nouns (27.3%) with a slightly lower precision of 93.9%. These results cannot be used successfully in the subsequent combination to complex semantic sorts. This is due to the fact that for each separate characteristic, many classified nouns are different, not producing enough overlap. Thus, the recall for some characteristics drops in both the lenient and the strict combination. The same accounts for the newly defined nouns and their binary characteristics.

Lastly, in case complex semantic sorts are required for the lexicon, the best results can be obtained by using the outcome of bootstrapping on verb-deep-subject relations as Table 6 indicates. Here, an average precision of 87.2% for all kinds of semantic sorts is still an improvement over previous methods.

5 Conclusion

This paper investigates the extension and improvement of the lexical acquisition approach as presented in (Biemann and Osswald, 2006). New relations such as the ones between a verb and its arguments have been utilized as input and have shown a significantly higher precision than modifying adjectives.

By using only nouns that were identified as having the same class in three different context types, we got a precision of 98.2% averaged over all binary semantic characteristics, with a recall of 14.7%. However, the maximally possible recall is bounded by the corpus and is at around 36% for the single context types.

We showed that by creating complex semantic sorts with the help of binary characteristics, new nouns from different kinds of sorts can be identified with a precision of about 87%. Finally, the high amount of single characteristics obtained for yet unknown nouns renders this approach very useful for lexical acquisition.

Acknowledgments

We would like to thank Sven Hartrumpf for helping us with parsing the corpus.

References

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of COLING/ACL'98*, Montréal, Canada.

Chris Biemann and Rainer Osswald. 2006. Automatic extension of feature-based semantic lexicons via contextual attributes. In *From Data and Information Analysis to Knowledge Engineering – Proceedings of the 29th Annual Conference of the Gesellschaft für Klassifikation*, pages 326–333. Springer.

Christian Biemann, Stefan Bordag, Gerhard Heyer, Uwe Quasthoff, and Christian Wolff. 2004. Language-independent methods for compiling monolingual lexical data. In *Proceedings of CicLING 2004*, LNCS 2945, pages 215–228, Berlin. Springer.

Xavier Carreras and Lluís Màrques. 2004. Introduction to the CoNLL-2004 shared task: Semantic role labeling. In *Proceedings of CoNLL-2004*, pages 89–97.

Masato Hagiwara, Yasuhiro Ogawa, and Katsuhiko Toyama. 2006. Selection of effective contextual information for automatic synonym acquisition. In *Proceedings of COLING/ACL 2006*, Sydney, Australia.

Zelig Harris. 1968. *Mathematical Structures of Language*. New York, Wiley.

Sven Hartrumpf, Hermann Helbig, and Rainer Osswald. 2003. The semantically based computer lexicon HaGenLex – Structure and technological environment. *Traitement automatique des langues*, 44(2):81–105.

Sven Hartrumpf. 2003. *Hybrid Disambiguation in Natural Language Analysis*. Der Andere Verlag, Osnabrück, Germany.

Sven Hartrumpf. 2005. Question answering using sentence parsing and semantic network matching. In *Multilingual Information Access for Text, Speech and Images: 5th Workshop of the Cross-Language Evaluation Forum, CLEF 2004*, LNCS 3491, pages 512–521. Springer.

Hermann Helbig. 2006. *Knowledge Representation and the Semantics of Natural Language*. Springer, Berlin.

Dekang Lin. 1998. Automatic retrieval and clustering of similar words. In *Proceedings of the 17th international conference on Computational linguistics*, pages 768–774.

Fernando Pereira, Naftali Tishby, and Lillian Lee. 1993. Distributional clustering of english words. In *Proceedings of the 31st annual meeting on Association for Computational Linguistics*, pages 183–190.

Ellen Riloff and Jessica Shepherd. 1997. A corpus-based approach for building semantic lexicons. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, pages 117–124.

Brian Roark and Eugene Charniak. 1998. Noun-phrase co-occurence statistics for semi-automatic semantic lexicon construction. In *Proceedings of COLING/ACL'98*, Montréal, Canada, pages 1110–1116.

Mats Rooth, Stefan Riezler, Detlef Prescher, Glenn Carroll, and Franz Beil. 1999. Inducing a semantically annotated lexicon via EM-based clustering. In *Proceedings of the 37th Annual Meeting of the ACL*.

Michael Thelen and Ellen Riloff. 2002. A bootstrapping method for learning semantic lexicons using extraction pattern contexts. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*.