Hybrid Disambiguation of Prepositional Phrase Attachment and Interpretation

Sven Hartrumpf
Applied Computer Science VII (AI)
University of Hagen
58084 Hagen, Germany
Sven.Hartrumpf@fernuni-hagen.de

Abstract

In this paper, a hybrid disambiguation method for the prepositional phrase (PP) attachment and interpretation problem is presented.[1] The data needed, semantic PP interpretation rules and an annotated corpus, is described first. Then the three major steps of the disambiguation method are explained. Cross-validated evaluation results for German (88.6–94.4% correct for binary attachment ambiguities, 83.3–92.5% correct for interpretation ambiguities) show that disambiguation methods combining interpretation rules and statistical methods might yield significantly better results than non-hybrid disambiguation methods.

1 Introduction

The problem of prepositional phrase (PP) attachment ambiguity is one of the most famous problems in natural language processing (NLP). In recent years, many statistical solutions have been proposed: lexical associations (see (Hindle and Rooth, 1993)); error-driven transformation learning (see (Brill and Resnik, 1994), extensions by (Yeh and Vilain, 1998)); backed-off estimation (see (Collins and Brooks, 1995), extended to the multiple PP attachment problem by (Merlo et al., 1997)); loglinear model (see (Franz, 1996b), (Franz, 1996a, pp. 97–108)); maximum entropy model (see (Ratnaparkhi, 1998; Ratnaparkhi et al., 1994)).

The disambiguation method in this paper has two key features. First, it tries to solve the PP attachment problem *and* the PP interpretation problem. Second, it is *hybrid* as it combines more traditional PP interpretation rules and statistical methods.

2 Data

2.1 PP interpretation rules

One central component for the disambiguation method presented in this paper are semantic interpretation rules for PPs. A PP interpretation rule consists of a premise and a conclusion. The premise of an interpretation rule describes under which conditions the PP interpretation specified by the rule's conclusion can be valid. Two example rules for the local and contents interpretation of 'über' ('about'/'above'/'on'/'over'/'via'/…) are shown in Figure 1. As (at least) five more interpretations of 'über' are possible, the ambiguity degree for the interpretation of such a PP is (at least) seven.

The premise of a rule is a set of feature structure constraints (including negated and disjunctive constraints and defining an underspecified feature structure) that refer to the following features of the preposition's sister NP (nominal phrase) and the preposition's mother NP or V (verb). (The features that are only refered to for the sister NP are marked by an S.)

case (S) syntactic case: genitive, dative, and accusative for German PPs

num (S) syntactic number: singular and plural in German

sort a semantic sort value (atomic or disjunctive value) from a predefined ontology (see (Helbig and Schulz, 1997)) comprising 45 sorts. The most important

[1]This disambiguation method was developed for an NLI in the *Virtuelle Wissensfabrik* (*Virtual Knowledge Factory*, see (Knoll et al., 1998)), a project funded by the German state Nordrhein-Westfalen, which supported this research in part. I would like to thank Rainer Osswald and the anonymous reviewers for their useful comments and suggestions.

id	über.loc
explanation	c1 is/happens above the location of c2.
examples	'Flugzeuge über Seen' ('air planes above lakes'), ...
premise	c1 (sort (dis object situation)) c2 (case dat) (sort concrete-object)
conclusion	net (loc c1 c3) (*ueber c3 c2)

id	über.mcont
explanation	c1 contains information about the topic described by c2.
examples	'Bücher über Seen' ('books on lakes'), ...
premise	c1 (sort (dis object situation)) (info +) c2 (case acc) (sort object)
conclusion	net (mcont c1 c2)

The semantic network node c1 corresponds to the mother, the node c2 to the sister, and c3 etc. are additional nodes. A disjunction of feature values is introduced by *dis*.

Figure 1: PP interpretation rules for two interpretations of 'über'

sorts for nouns are *object* and its subsorts *con-object* (concrete object, with subsorts *dis-object* (discrete object) and *substance*) and *abs-object* (abstract object, with subsorts *tem-abstractum* (temporal abstractum), *abs-situation* (abstract situation), *attribute*, etc.). Verbs can belong to sort *stat-situation* (static situation) or sort *dyn-situation* (dynamic situation, with subsorts *action* and *event*). A disjunctive value represents a *concept family* (as introduced by (Bierwisch, 1983); closely related are *dotted types*, see for example (Buitelaar, 1998)), e.g., the noun 'book' comprises a physical object *variant* and an abstract information *variant*.

etype extension type for distinguishing individuals ('child', 'table'), sets of individuals ('men', 'group', 'people'), etc.

The rest of the features are semantic Boolean features as shown in Table 1.[2]

The conclusion of a rule is a semantic interpretation of the PP, which can be valid if the premise is satisfied by the sister and the mother. The rules' semantic representation uses a multilayered extended semantic network formalism (MESNET, see for example (Helbig and Schulz, 1997)), which has been successfully applied in various areas (e.g., in the *Virtual Knowledge Factory*, see (Knoll et al., 1998)).

Besides the premise and the conclusion, each rule contains a mnemonic identifier like *in.loc* (which consists of the preposition's orthographic form followed by an abbreviation derived from the semantic interpretation in the conclusion), a short explanation, and a set of example sentences that can be interpreted using this rule.

From a set of rules for 160 German prepositions collected by (Tjaden, 1996), all rules for six important (i.e., frequent) prepositions were taken as a starting point for development and evaluation of a hybrid disambiguation method. Sentences were retrieved from a development test corpus to refine these rules.

2.2 Corpus

While PP interpretation rules form the **rule component** of the hybrid disambiguation method, an annotated corpus serves as the source of the **statistical component**. For each preposition under investigation, a number of candidate sentences that possibly show attachment ambiguity for this preposition were automatically extracted from a corpus. This corpus is based on the online version of the Süddeutsche Zeitung, starting from August 1997. The corpus is marked up according to the Corpus Encoding Standard (see (Ide et al., 1996)) and word, sentence, and paragraph identifiers are assigned.

The preposition in a candidate sentence is semiautomatically annotated with five attributes:

sister The position of the right-most word of the preposition's sister NP. Postnominal genitive NPs modifying the main sister NP are included in this annotation.

[2] Of course, other sets of such features are possible; the choice was made by selecting relevant features from the set of semantic features in an existent German inheritance lexicon (see (Hartrumpf and Schulz, 1997)), which contains 7000 lexemes and is used by the disambiguation method.

feature name	description of entities with positive (+) value	examples
animate (S)	an animate entity	'animal', 'person', 'tree'
geogr	a geographical concept	'city', 'country'
human	a human entity	'child', 'president'
info	an entity that carries information	'book', 'concert'
instit	an institution	'company', 'parliament'
instru (S)	an entity that can be used as an instrument	'hammer', 'ladder'
legper	a legal person	'company', 'woman'
mental	a mental state or process	'fear', 'happiness'
method	a method	'compression', 'filtering'
potag	a (potential) agent	'horse', 'man'

Table 1: Semantic Boolean features in PP interpretation rules

mother The position of the syntactic head word of the mother NP or V.

amother The list of **alternative mothers** represented by the position of the syntactic head word of an NP or V. An alternative mother is a syntactically possible mother distinct from the (correct) mother. All alternative mothers plus the (correct) mother form the set of **candidate mothers** for PP attachment.

c-id A character string that identifies the semantic reading of the preposition and corresponds to the identifier in a PP interpretation rule (see Figure 1).

c A character string for comments and documentation purposes.

The preposition in corpus sentence (1) is annotated as shown by the SGML element in (2). The meaning of this annotation can be illustrated as in (3): the PP's sister ends at 'Seite'; the PP attaches to 'gebaut', and could syntactically also be attached to the NP with head 'Depot' or the NP with head 'Museums'; the interpretation of the PP is a local one (*auf.loc*).[3]

(1) *Und wieso wird das neue Depot*
And why is the new depot
des Deutsch-Deutschen
the+GEN German-German
Museums auf bayerischer Seite
Museum on Bavarian side
gebaut, nachdem die Planungen für
built, after the plannings for
die Thüringer Talseite schon
the Thuringian valley-side already
fertig waren?
ready were?
'And why is the new depot of the German-German Museum built on the Bavarian side, after the planning for the Thuringian side of the valley has already been completed?'

(2) 19971002bay_c.p3.s2.w10 (article bay_c, 1997-10-02, paragraph 3, sentence 2, word 10): ⟨w c-id="auf.loc" sister="12" mother="13" amother="6/9"⟩auf⟨/w⟩

(3) Und wieso wird das neue **Depot**^{a1} des Deutsch-Deutschen **Museums**^{a2} **auf**$^{auf.loc}$ bayerischer **Seite**^{s} **gebaut**^{m}, nachdem die Planungen für die Thüringer Talseite schon fertig waren?

The annotation process is semiautomatic: the machine guesses the attribute values following some heuristics; these guesses have to be checked and possibly extended or corrected by a human annotator. This kind of annotation, of course, is labor-intensive. But due to the development of an Tcl/Tk annotation tool optimized for manual annotation speed, the average annotation time per candidate sentence dropped under 30 seconds. Furthermore, the following sections show that a small set of annotated sentences achieves promising results for PP attachment and interpretation. The lexicon (see footnote 2) had to be extended for the nouns and

[3]Please note that the translations of sentences (1) and (4) are not ambiguous.

verbs annotated as head words of sisters or candidate mothers that were not in the lexicon and could not be analyzed by a compound analysis module.

Some candidate sentences were excluded from the investigation because the PP involves a problem that is supposed to be solved by other NLP modules[4] and could disturb the evaluation of the PP disambiguation module (e. g., by producing noise for the statistical part). All exclusion criteria are listed in Table 2 with percentages of instances of such exclusions relative to the number of candidate sentences. In short, sentences are excluded when their PP ambiguity problem

- can be solved by separate components (for support verb constructions and idioms) or
- can only be solved if the PP attachment and interpretation is supported by another component (for complex named entities, ellipsis resolution, and foreign language expressions).

The first 120 non-excluded candidate sentences for each preposition were chosen and randomally split into eight parts for *cross validation*. Eight evaluations were carried out with one part being the evaluation test corpus and the remaining seven parts being the evaluation training corpus.

Sometimes, it makes no semantic difference whether a PP in a sentence attaches to an NP or a V. This is known as *systematic ambiguity* (or *systematic indeterminacy*, see (Hindle and Rooth, 1993, p. 112)). Two subtypes of this phenomenon are systematic locative ambiguity (see corpus sentence (4)) and systematic contents ambiguity.

(4) *Bis* *ein* ***Bescheid***^{m1} ***aus***$^{aus.origl}$
Until a notification from
Karlsruhe^{s} ***eintrifft***^{m2}, *kann* *es*
Karlsruhe comes-in, can it
Monate *dauern.*
months take.
(19971001fern_d.p3.s6.w4)

'It might take months until a notification from Karlsruhe comes in.'

[4]It should be evaluated in further research how well such modules solve these problems.

The frequency of such ambiguities depends heavily on the preposition; on the average, there were 4.3% cases of systematic ambiguity.[5] For English, (Hindle and Rooth, 1993, p. 116) report that 77 out of 880 sentences (8.75%) were systematically ambiguous. In such sentences, an attachment can be considered correct if it is one of the two attachments connected by systematic ambiguity; both parsing results will lead to identical results in an NLP application if it contains sufficiently developed inference components. Table 3 shows for the evaluation corpus (720 sentences[6]) where the PP attaches to (columns V, NP1, NP2 (the second closest NP), NP3, NP4), how many attachments are syntactically possible (number of candidate mothers; columns labeled 1 to 5), and how frequent systematic ambiguity is (last column).

3 Hybrid disambiguation method

3.1 Basic ideas

PP attachment is one of the most famous problems in NLP. But where a PP *attaches* to, is only half of the story of the PP's contribution to an utterance; the other half is how it is to be *interpreted.* And clearly, these two questions are not independent. So, why not tackle both problems *at once*, trying to achieve for both problems results that are better than the results obtained by an isolated PP attachment component and an isolated PP interpretation component? As both problems depend on each other, there is the strong hope that this is the case. To investigate this hypothesis, such a disambiguation method was developed and evaluated.

The input to the disambiguation method is the feature structure p for the preposition, the feature structure s for the parse of the preposition's sister NP, and the feature structures cm_i for the (trivial) parses of the syntactic head words of all candidate mothers. The output is the mother the PP is to be attached to and the interpretation the preposition plus the sister NP contribute to the meaning of the enclosing sentence.

The overall structure of this disambiguation method comprises three steps. First, all sets

[5]All annotated sentences showing systematic ambiguity contain only the two candidate mothers that are related by the underlying systematic ambiguity.

[6]These annotated sentences are available for research.

short name	description	% of tokens
cne-amother	amother is a complex named entity (titles of books, etc.)	0.1
cne-mother	mother is a complex named entity (titles of books, etc.)	0.4
cne-sister	sister is a complex named entity (titles of books, etc.)	0.6
ell-amother	amother is elliptic	0.1
ell-mother	mother is elliptic	0.1
ell-sister	sister is elliptic	0.5
fle-amother	amother is a foreign language expression	0.1
fle-mother	mother is a foreign language expression	0.1
idi-amother	amother is an idiom (or part of an idiom)	0.1
idi-mother	mother is an idiom	0.4
idi-pp	PP is an idiom	3.6
idi-pp-mother	PP plus mother is an idiom	0.9
idi-pp-v	PP plus verb is an idiom	0.5
problem	unclassified problem	0.7
svc	PP is part of a support verb construction	0.5
svc-amother	amother of the PP is a support verb construction	0.3
svc-mother	mother of the PP is a support verb construction	1.0
sum		10.1

Table 2: Exclusion criteria for candidate sentences

preposition	observed attachment %					ambiguity degree %					sys. amb. %
	V	NP1	NP2	NP3	NP4	1	2	3	4	5	
auf	56.7	38.3	5.0	0.0	0.0	13.3	58.3	24.2	2.5	1.7	5.0
aus	22.5	75.0	2.5	0.0	0.0	35.8	51.7	8.3	4.2	0.0	10.0
bei	52.5	42.5	5.0	0.0	0.0	30.8	51.7	14.2	1.7	1.7	6.7
über	37.1	57.1	5.0	0.8	0.0	17.5	66.7	13.3	0.8	1.7	2.5
vor	41.3	52.1	5.0	1.7	0.0	23.3	61.7	13.3	1.6	0.0	0.8
wegen	62.1	26.3	10.0	1.7	0.0	9.2	74.2	14.2	1.7	0.8	0.8
average	45.4	48.5	5.4	0.7	0.0	21.7	60.7	14.6	2.1	1.0	4.3

Table 3: Attachment data from the evaluation corpus

of possible interpretations PI_i of the PP plus a given candidate mother cm_i are determined by applying the PP interpretation rules. Second, for each set of possible interpretations PI_i, one interpretation si_i is selected using interpretation statistics (on semantics). Third, among all selected si_i, one interpretation is chosen based on attachment statistics (on semantics and syntax) and additional factors. These steps will be presented in more detail in the following three subsections.

3.2 Application of interpretation rules

Step 1 of the disambiguation method (determining possible interpretations PI_i) is driven by testing the premises of PP interpretation rules. From the set of interpretations PI'_i whose rule premises are satisfied, interpretations are removed that violate adjunct constraints from the lexicon or constraints from the underlying semantic formalism[7] (see step 1 in Figure 2).

[7]Of course, constraints from the semantic formalism could be added to the rules. But this would introduce redundancy which would make the rules difficult to develop and maintain.

n is the number of possible attachments ($cm_1, \ldots, cm_n$).
m is the number of rules for preposition p ($r_1, \ldots, r_m$).

1. for each candidate mother cm_i
   (a) $PI'_i = \{(p, s, cm_i, r_j) \mid 1 \leq j \leq m, \text{premise of rule } r_j \text{ is satisfied by sister } s \text{ and } cm_i\}$
   (b) PI_i = set of all $(p, s, cm_i, r) \in PI'_i$ which fulfill the following conditions:
      - Semantic relations in the conclusion of r are licensed by compatible relations listed in the feature structure cm_i, which come from lexical entries (or lexical defaults).
      - Semantic relations in the conclusion of r do not violate the signature constraints that are defined for these relations in the underlying semantic network formalism.
2. for each candidate mother cm_i with nonempty PI_i
   (a) $si_i = \arg\max_{pi} rf(r, \{r_j | \exists(p, s, cm_i, r_j) \in PI_i\})$, where $pi = (p, s, cm_i, r) \in PI_i$
3. for each candidate mother cm_i with nonempty PI_i
   (a) d = distance in words between candidate mother cm_i and the PP (p plus s)
   (b) $score_{si_i} = rf((r, cat(cm_i)), \{(r_j, cat(cm_k)) \mid 1 \leq k \leq n, PI_k \neq \emptyset, si_k = (p, s, cm_k, r_j)\})$ $+ \; score_{dist}(d)$, where $si_i = (p, s, cm_i, r)$

si = $\arg\max_{si_i} score_{si_i}$, where $1 \leq i \leq n, PI_i \neq \emptyset$

Figure 2: Disambiguation algorithm

To simplify Figure 2, the treatment of complements is excluded. Interpretations that are licensed by lexical complement information for candidate mothers are also determined in step 1. Experiments showed that it is a good strategy to prefer complement interpretations over adjunct interpretations, which are described in the following steps.[8] Attachment cases where prepositional objects as complements are involved are the easy ones for statistical disambiguation techniques (see for example (Hindle and Rooth, 1993)); in a hybrid system, one can expect such complement information to be in the lexicon, at least in part. The problem is alleviated as the interpretation rules (which are developed for adjuncts) produce correct results for many complements; but this topic needs further research.

3.3 Interpretation disambiguation

The result of step 1 can be viewed as an attachment-interpretation matrix ($ai_{i,j}$) with size $n \times m$. A matrix element $ai_{i,j}$ corresponds to attaching the PP to candidate mother cm_i under interpretation r_j and represents some kind of preference score.

To solve the attachment and interpretation problem (i. e., to select the right matrix element), statistics can be used. There are numerous statistical approaches (see section 1), but in the presented approach a statistical component is combined with a rule component (see step 1). This rule component reduces the degree of ambiguity (i. e., marks elements in matrix ($ai_{i,j}$) as possible or impossible) and delivers high-level semantic information (the possible semantic interpretations of the PP for a given candidate mother) for statistical disambiguation.

The strategy adopted in this disambiguation method is to do the remaining disambiguation in two steps: first disambiguate the interpretations for each attachment possibility, then disambiguate the attachments based on the first step's result. So, in step 2 of the disambiguation method, one interpretation for each candidate mother is chosen. As Table 4 shows, most of the time the correct rule fires (given the correct mother; see recall column), but false rules fire too (see precision column) because interpretation rules refer only to a limited depth

[8] In the rare case of two possible complement interpretations, the verbal one is prefered.

preposition	readings	recall %	precision %
auf	9	100.0	100.0
aus	6	97.4	39.8
bei	4	93.7	69.8
über	7	100.0	65.4
vor	6	98.3	54.7
wegen	1	100.0	100.0

Table 4: Results of PP interpretation rules for (correct) mothers

$$rf(\text{aus.pars}, \{\text{aus.origl}, \text{aus.pars}, \text{aus.sourc}\}) = 1.0$$
$$rf((\text{aus.temp}, \text{np}), \{(\text{aus.cstr}, \text{v}), (\text{aus.temp}, \text{np})\}) = 1.0$$

Figure 3: Statistical example data for interpretation and attachment

of semantics, which can be delivered by realistic parsers for nontrivial domains. Therefore, there is the need to disambiguate for interpretation. Here statistics derived from the annotated corpus come into play: relative frequencies are calculated, which serve as estimated probabilities.

As usual in statistical methods for disambiguation, there is a trade-off between depth of learned information (e. g., number and type of features) and non-sparseness of the resulting matrix-like structure representing the learning results: the deeper the information, the sparser the matrix. A good compromise for the problem at hand is to regard only the interpretation (identified by the rule id) and to establish a limit n_{int} for the number of interpretations. Empirical results showed that three is a reasonable choice for n_{int}. An example of an entry in the interpretation statistics is given in the first line of Figure 3 and can be paraphrased as follows: The interpretation *aus.pars* wins in 100% of the learned cases if the interpretations *aus.origl* and *aus.sourc* are possible too.

If there are more than three possible interpretations, standard techniques for reducing to several triples can be used (*backed-off estimation*, see for example (Katz, 1987), (Collins and Brooks, 1995)). The relative frequency of rule r_i being the correct interpretation among $I = \{r_1, r_2, \ldots, r_n\}$ is estimated for $n > n_{int}$ as in equation (5):

$$(5) \quad rf(r_i, I) := \frac{\sum_{c \in C_i} rf(r_i, c)}{|C_i|}$$

where C_i is the set of all subsets of I with n_{int} elements that contain r_i.

In step 2 of the disambiguation algorithm (see middle of Figure 2), the rule that maximizes the (estimated) relative frequency must be found for each candidate mother.

3.4 Attachment disambiguation

After step 2, the attachment-interpretation matrix $(ai_{i,j})$ contains in each row (attachment) one element marked as selected.[9] What remains to be done is to choose among all attachments with selected interpretation si_i one interpretation si.

For this disambiguation task, attachment statistics are employed. This time the compromise between depth of learned information and non-sparseness can contain more information than just the interpretation id as experiments showed. A three-valued syntactic-semantic feature *cat* is added. It describes the candidate mother with three possible values:

v a verb

nps an NP that describes a situation (at least partially), e. g., 'continuation'

np an NP that does not describe a situation, e. g., 'house'

The second line of Figure 3 contains an example that expresses the fact that if the interpretation *aus.temp* for a nominal candidate mother and the interpretation *aus.cstr* for a verbal candidate mother compete then the first is correct (in the training corpus) with relative frequency 1. If one adds even more information to attachment statistics (e. g., the position of NP candidate mothers like np2 for the second closest NP) the attachment data for the annotations in this paper becomes too sparse.

[9] There might be rows where no element is marked because none of the rules fired and passed filtering (see section 3.2).

As for the interpretation statistics in step 2, standard techniques can reduce tuples that are longer than 2 (n_{att}) to several shorter ones. The relative frequency of $(r_i, cat(cm_i))$ belonging to the correct attachment among $A = \{(r_1, cat(cm_1)), \ldots, (r_n, cat(cm_n))\}$ is estimated for $n > n_{att}$ as in equation (6):

(6) $$rf((r_i, cat(cm_i)), A) := \frac{\sum_{c \in C_i} rf((r_i, cat(cm_i)), c)}{|C_i|}$$

where C_i is the set of all subsets of A with n_{att} elements that contain $(r_i, cat(cm_i))$.

These relative frequencies for the selected interpretations si_i serve as initial values for an *attachment score*. Other factors can add to this score, so that the attachment decision should improve; of course, the value is only a score, not a relative frequency any more. Different factors (e. g., distance between candidate mother and the PP; in this way, one can simulate the right-association principle, see (Kimball, 1973)) were evaluated. The following distance scoring function $score_{dist}$ turned out to be useful:

(7) d is the number of words between the candidate mother and the PP. m_d is an upper limit for distances. Longer distances are reduced to m_d. (10 is a reasonable choice for m_d.)

$$score_{dist}(d) := \begin{cases} \frac{dist_w \cdot (m_d - \min(d, m_d))}{m_d} & \text{for NP mothers} \\ \frac{dist_w \cdot (m_d - \min(d \cdot dist_v, m_d))}{m_d} & \text{for V mothers} \end{cases}$$

Good values for the parameters $dist_w$ (weight of the distance factor) and $dist_v$ (modification for verbal mothers) depend on the preposition at hand and are learned by testing pairs of values from the range 0.0 to 2.0 (see Table 5).[10] The last step of the disambiguation algorithm is summarized at the bottom of Figure 2.

preposition	$dist_w$	$dist_v$
auf, vor, wegen	0.8	0.6
aus	1.2	1.0
bei	1.2	0.8
über	0.8	0.2

Table 5: Good parameters for the attachment scoring function $score_{dist}$

4 Evaluation

Cross validation (see section 2.2) showed that hybrid disambiguation achieves for both problems, PP attachment and PP interpretation ambiguity, satisfying correctness results for all six prepositions (see Table 6): 88.6–94.4% for binary attachment ambiguities, 85.6–90.8% for all ambiguous attachments, and 75.0–84.2% for ambiguity degrees above 2 (leading to the multiple PP attachment problem).

Comparison of the interpretation results is impossible as these are the first cross-validated results for PP interpretation. But 83.3–92.5% correctness for prepositions with more than one reading seems very promising.

Comparison of the attachment results is possible, but difficult. One reason is that the best reported disambiguation results for binary PP attachment ambiguities (84.5%, (Collins and Brooks, 1995); 88.0% using a semantic dictionary, (Stetina and Nagao, 1997)) are for English. Because word order is freer in German than in English, the frequency and degree of attachment ambiguity is probably higher in German. There are only few evaluation results for German: (Mehl et al., 1998) achieve 73.9% correctness for the preposition 'mit' ('with'/'to'/...) using a statistical lexical association method.

Of course, the evaluation corpus is not large (720 sentences); so, the results reported in this paper must be treated with some caution. But as the selected prepositions show diverse numbers of readings (1–9, see Table 4) and the results are cross-validated, it is likely that the reported results will not deteriorate for larger corpora.

5 Conclusions

In this paper, a new hybrid disambiguation method which uses PP interpretation rules and

[10] The best values for these parameters probably also depend on text type, text domain, interpretation of the PP, etc.

preposition	correctness in percentage								
	attachment for ambiguity degree							interpretation	att. and int.
	1	2	3	4	5	≥2	≥3		
auf	100.0	88.6	75.9	100.0	100.0	85.6	79.4	92.5	86.7
aus	100.0	90.3	80.0	80.0	–	88.3	80.0	90.8	85.8
bei	100.0	90.3	82.4	50.0	50.0	86.7	76.2	91.7	85.0
über	100.0	88.8	81.3	100.0	100.0	87.9	84.2	83.3	83.3
vor	100.0	89.2	75.0	100.0	–	87.0	77.8	89.2	81.7
wegen	100.0	94.4	70.6	100.0	100.0	90.8	75.0	100.0	91.7

Table 6: Results of hybrid disambiguation

statistics about attachment and interpretation in an annotated corpus was described. It yields results with competitive correctness for both the PP attachment problem and the PP interpretation problem.

Some questions had to be left open, e.g., a nontrivial reading disambiguation[11] for candidate mothers and sister NPs. Questions concerning the requisite manual work (maintaining rules and some parts of annotating corpora) arise: How much does this work pay off and how could more of this work be automated? The disambiguation method should be evaluated for larger corpora (more sentences, more prepositions) in future research. The ongoing use of the disambiguation method in natural language interfaces will provide valuable feedback.

References

Manfred Bierwisch. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In Rudolf Ružička and Wolfgang Motsch, editors, *Untersuchungen zur Semantik*, Studia grammatica XXII, pages 61–99. Akademie-Verlag, Berlin.

Eric Brill and Philip Resnik. 1994. A rule-based approach to prepositional phrase attachment disambiguation. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING 94)*, pages 1198–1204.

Paul Buitelaar. 1998. *CoreLex: Systematic Polysemy and Underspecification.* PhD dissertation, Brandeis University.

[11]This is closely related to the problem of word sense disambiguation; currently, this disambiguation is based on frequencies.

Michael Collins and James Brooks. 1995. Prepositional phrase attachment through a backed-off model. In *Proceedings of the 3rd Workshop on Very Large Corpora (WVLC-3).*

Alexander Franz. 1996a. *Automatic Ambiguity Resolution in Natural Language Processing*, volume 1171 of *LNAI.* Springer, Berlin.

Alexander Franz. 1996b. Learning PP attachment from corpus statistics. In Stefan Wermter, Ellen Riloff, and Gabriele Scheler, editors, *Connectionist, Statistical, and Symbolic Approaches to Learning for Natural Language Processing*, volume 1040 of *LNAI*, pages 188–202. Springer, Berlin.

Sven Hartrumpf and Marion Schulz. 1997. Reducing lexical redundancy by augmenting conceptual knowledge. In Gerhard Brewka, Christopher Habel, and Bernhard Nebel, editors, *Proceedings of the 21st Annual German Conference on Artificial Intelligence (KI-97)*, number 1303 in Lecture Notes in Computer Science, pages 393–396, Berlin. Springer.

Hermann Helbig and Marion Schulz. 1997. Knowledge representation with MESNET: A multilayered extended semantic network. In *Proceedings of the AAAI Spring Symposium on Ontological Engineering*, pages 64–72, Stanford, California.

Donald Hindle and Mats Rooth. 1993. Structural ambiguity and lexical relations. *Computational Linguistics*, 19(1):103–120, March.

Nancy Ide, Greg Priest-Dorman, and Jean Véronis, 1996. *Corpus Encoding Standard.* http://www.cs.vassar.edu/CES/.

Slava M. Katz. 1987. Estimation of probabilities from sparse data for the language model component of a speech recognizer. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-35(3):400–401, March.

John Kimball. 1973. Seven principles of surface structure parsing in natural language. *Cognition*, 2:15–47.

A. Knoll, C. Altenschmidt, J. Biskup, H.-M. Blüthgen, I. Glöckner, S. Hartrumpf, H. Helbig, C. Henning, Y. Karabulut, R. Lüling, B. Monien, T. Noll, and N. Sensen. 1998. An integrated approach to semantic evaluation and content-based retrieval of multimedia documents. In C. Nikolaou and C. Stephanidis, editors, *Proceedings of the 2nd European Conference on Digital Libraries (ECDL'98)*, volume 1513 of *LNCS*, pages 409–428, Berlin. Springer.

Stephan Mehl, Hagen Langer, and Martin Volk. 1998. Statistische Verfahren zur Zuordnung von Präpositionalphrasen. In Bernhard Schröder, Winfried Lenders, Wolfgang Hess, and Thomas Portele, editors, *Proceedings of the 4th Conference on Natural Language Processing - KONVENS-98*, number 1 in Computers, Linguistics, and Phonetics between Language and Speech, pages 97–110, Frankfurt, Germany. Peter Lang.

Paola Merlo, Matthew W. Crocker, and Cathy Berthouzoz. 1997. Attaching multiple prepositional phrases: Generalized backed-off estimation. In *Proceedings of the 2nd Conference on Empirical Methods in Natural Language Processing (EMNLP-2)*, pages 149–155, Providence, Rhode Island. Brown University.

Adwait Ratnaparkhi, Jeff Reynar, and Salim Roukos. 1994. A maximum entropy model for prepositional phrase attachment. In *Proceedings of the ARPA Human Language Technology Workshop*, pages 250–255.

Adwait Ratnaparkhi. 1998. Statistical models for unsupervised prepositional phrase attachment. In *Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics (COLING-ACL'98)*, pages 1079–1085.

Jiri Stetina and Makoto Nagao. 1997. Corpus-based PP attachment ambiguity resolution with a semantic dictionary. In *Proceedings of the 5th Workshop on Very Large Corpora (WVLC-5)*, pages 66–80.

Ingo Tjaden. 1996. Semantische Präpositionsinterpretation im Rahmen der Wortklassengesteuerten Analyse. Master's thesis, Fern-Universität Hagen, Hagen.

Alexander S. Yeh and Marc B. Vilain. 1998. Some properties of preposition and subordinate conjunction attachments. In *Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics (COLING-ACL'98)*, pages 1436–1442.