Using Tri-lexical Dependencies in LFG Parse Disambiguation

Aoife Cahill, Uli Heid, Christian Rohrer and Marion Weller
IMS, Universität Stuttgart, Germany

1   Introduction

The use of lexical dependencies in parse disambiguation is not a new idea. For example, it is one of the key ideas behind the Collins (1999) parser. In that parser, bi-lexical dependencies are used, but in later work Bikel (2004) showed that the bi-lexical dependencies, in fact had little or no impact on the parser decisions. One of the reasons these bi-lexical dependencies are thought to be ineffective, is because of sparse data. In treebank-derived parsers, there are simply not enough instances of the dependencies for them to have a significant impact.

Intuitively, on the other hand, it seems that including lexical dependencies should help. In this paper, we consider tri-lexical dependencies between verbs, prepositions and nouns, automatically extracted from a very large corpus, independent from any training data the parser uses. Consider the following example from German with the trilexical dependency between the verb stehen (to stand), the preposition zu (to), and the noun Verfügung (disposition), which correspond to the dependency *to be available* in English:

(1)   Dass das Geld   zur   Verfügung steht,   wird angenommen.
       That the money to the disposition stands, is   assumed.

       'The fact that the money is available, is assumed.'

For a parser, there is ambiguity here about where to attach the PP *zur Verfügung*, either to the DP *das Geld* or to the verb *stehen*. We observe that the PP almost never attaches to the DP.[1]

In this paper we attempt to incorporate this observation (and test its validity) in an LFG parse disambiguation scenario. Here we present preliminary results.

2   Tri-lexical Dependencies

We automatically acquire very large lists of tri-lexical dependencies from more than 234 million tokens of German newspaper text. The text is first parsed using the `fspar` parser (Schiehlen, 2003). In each sentence, the finite verbs are selected, and we check whether there is a PP (possibly) attached to this verb.[2] Reflexive verbs are differentiated from non-reflexives. We take note of the preposition and the head word of the prepositional object, giving us a tri-lexical dependency between a verb, preposition and noun. We use the UCS Toolkit (Evert, 2004) to calculate the association measures of the triples using log likelihood. Since the log likelihood is a two-place measure, and we have three elements, we calculate two types of association measures. The first calculates the association between the preposition+verb and the noun, and the other calculates the association between the preposition+noun and the verb.

3   Parsing System

In our experiments we use the handcrafted German LFG of Rohrer and Forst (2006) coupled with a log-linear disambiguation component (Riezler et al., 2002; Forst, 2007). This is a robust large-scale grammar that has been implemented within the XLE system and achieves completes parses for around 80% of newspaper text.

For both training and evaluation of the log-linear disambiguation models described in this paper, we use data constructed with the help of the TIGER Treebank (Brants et al., 2002). Our training data consists of 5617 pairs of labeled and unlabeled packed representations of c- and f-structures which have been produced by our grammar. The labeled representations were constructed by matching the f-structure part of the unlabeled representations produced by the grammar against packed f-structure representations that were derived from the original TIGER Treebank graphs (Forst, 2003). Only sentences for which a proper subset of the readings is compatible with the treebank annotations (and can be determined as such in a reasonable amount of time) were included in the training data, since only these are useful for discriminative training. For evaluation, we use the TiGer Dependency Bank (TiGer DB) (Forst et al., 2004), a dependency-based gold standard for German parsers.

4   Preliminary Experiments

We carry out a number of experiments to test the effect of tri-lexical dependencies in parse disambiguation. In all experiments, we only take the 40,000 dependencies with the highest log-likelihood ratios. On manual random inspection, the quality and reliability of the dependencies extracted decreases as the log likelihood decreases. This number was chosen arbitrarily and further experiments will be carried out to search for the ideal number of dependencies.

---

[1]One exception are PPs headed by *von* which do not behave regularly in their attachment.

[2]Incorrect PP attachments from the parser will obviously lead to incorrect dependencies, but since we sort by frequency at the end, we hope that the noise gets filtered out.

In each of the training sentences (both labeled and unlabeled), if a verb, preposition, noun dependency is found (where the verb (indirectly) dominates the PP), we add a feature to the f-structure noting whether the PP attaches to the DP or higher up at the verbal level. We experiment with five different models: (1) a model including the log-likelihood of the preposition+verb and noun, and a feature whether the PP attaches to DP or to the verbal level, (2) a model including the log-likelihood of the preposition+noun and verb, and a feature whether the PP attaches to DP or to the verbal level, (3) a model including the log-likelihood of the preposition-verb and a feature for *each* preposition whether the PP attaches to DP or to the verbal level (e.g. DP_ATTACH_nach), (4) whether accusative PPs (in general) attach to DP or to the verbal level[3], (5) a combination of (1) - (4).

We evaluate each model on our test set of 1430 TiGer DB sentences. The results as compared to a baseline model with only the features described in Forst (2007) are presented in Table 1. We see that not all models out-perform the baseline, and in fact the combination of all models yields a (significantly) lower result. The model which includes the log-likelihood of the verb+preposition as a feature performs much better than the model with the log-likelihood of the noun+preposition. The differences between the baseline and models (1) and (5) is found to be statistically significant ($p < 0.05$) using the Approximate Randomization Test. All other results are not statistically significantly different from the baseline.

Model	Baseline	1	2	3	4	5
Most-probable f-score	80.92	81.04	80.89	80.99	81.05	80.71

Table 1: Preliminary Results

An example of a sentence, where parse disambiguation improves with the new features in all 5 models is given in (2). In this case, the system has learned that the PP *am Dienstag* (on Tuesday) does not attach to the preceding DP, but rather at the verbal level dominated by the verb *entscheiden* (to decide).

(2) Das | entschied | das Bundesverfassungsgericht   (BVG) | am || Dienstag |.
     That   decided   the Federal Constitutional Court (BVG) on   Tuesday.

     'The Federal Constitutional Court decided that on Tuesday.'

5 Conclusion

We found that in a preliminary study of including tri-lexical dependencies into an LFG parse disambiguation model, overall disambiguation results can be improved.

References

Bikel, D. M. (2004). A distributional analysis of a lexicalized statistical parsing model. In *In EMNLP*, pages 182–189.

Brants, S., Dipper, S., Hansen, S., Lezius, W., and Smith, G. (2002). The TIGER treebank. In *Proceedings of the Workshop on Treebanks and Linguistic Theories*, Sozopol, Bulgaria.

Collins, M. (1999). *Head-Driven Statistical Models for Natural Language Parsing*. PhD thesis, MIT.

Evert, S. (2004). *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. PhD thesis, University of Stuttgart.

Forst, M. (2003). Treebank Conversion – Creating a German f-structure bank from the TIGER Corpus. In *Proceedings of the LFG03 Conference*, Saratoga Springs, NY. CSLI Publications.

Forst, M. (2007). *Disambiguation for a Linguistically Precise German Parser*. PhD thesis, University of Stuttgart.

Forst, M., Bertomeu, N., Crysmann, B., Fouvry, F., Hansen-Schirra, S., and Kordoni, V. (2004). Towards a dependency-based gold standard for German parsers – The TiGer Dependency Bank. In *Proceedings of the COLING Workshop on Linguistically Interpreted Corpora (LINC '04)*, Geneva, Switzerland.

Riezler, S., King, T. H., Kaplan, R. M., Crouch, R., Maxwell, J. T., and Johnson, M. (2002). Parsing the Wall Street Journal using a Lexical-Functional Grammar and Discriminative Estimation Techniques. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics 2002*, Philadelphia, PA.

Rohrer, C. and Forst, M. (2006). Improving coverage and parsing quality of a large-scale LFG for German. In *Proceedings of the Language Resources and Evaluation Conference (LREC-2006)*, Genoa, Italy.

Schiehlen, M. (2003). A Cascaded Finite-State Parser for German. In *Proceedings of the Research Note Sessions of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL'03)*, Budapest.

---

[3]Accusative PPs in most cases are directional, even in the rare cases where they attach to nouns, e.g. die Brücke über den Fluss (the bridge over the river).