Ungreedy Methods for Chinese Deterministic Dependency Parsing

Xiangyu Duan, Jun Zhao, Bo Xu

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
Zhongguancun Eastern Road 95#, 100080, Beijing, China
xyduan@nlpr.ia.ac.cn, jzhao@nlpr.ia.ac.cn, xubo@hitic.ia.ac.cn

Abstract

Deterministic dependency parsing has often been regarded as an efficient parsing algorithm while its parsing accuracy is a little lower than the best results reported by more complex parsing models. In this paper, we compare deterministic dependency parsers with complex parsing methods such as generative and discriminative parsers on the standard data set of Penn Chinese Treebank. The results show that, for Chinese dependency parsing, deterministic parsers outperform generative and discriminative parsers. Furthermore, based on the observation that deterministic parsing algorithms are greedy algorithms which choose the most probable parsing action at every step, we propose three kinds of ungreedy deterministic dependency parsing algorithms to globally model parsing actions. Results show that ungreedy deterministic dependency parsers perform better than original deterministic dependency parsers while maintaining the same time complexity, and our best parser improves much over all other parsers.

Introduction

Syntactic parsing is one of the most important tasks in Natural Language Processing (NLP). The mainstream of syntactic parsing is the statistical method that often focuses on generative and discriminative models. These models perform well while the time complexity is very high. Deterministic parsers emerge as efficient algorithms that take parsing actions stepwisely on the input sentence, and reduce the time complexity to linear or quadratic with the sentence's length. Deterministic parsers were firstly proposed for dependency parsing (Yamada and Matsumoto, 2003; Nivre and Scholz, 2004). Later, Sagae and Lavie (2005) and Wang et al. (2006) applied deterministic parsing for phrase structure parsing.

On the standard data set of Penn English Treebank, deterministic parsers show great efficiency in terms of time, offering accuracy just below the state-of-the-art parsing methods for English. In this paper, for Chinese dependency parsing, we use algorithms of Yamada and Matsumoto (2003), Nivre and Scholz (2004), and do the comparison of deterministic parsers with generative and discriminative parsers on Penn Chinese Treebank version 5.0. The results show that deterministic parsers perform much better than generative and discriminative parsers.

Furthermore, we observe that deterministic parsers are greedy. They choose the most probable parsing action at every step and loose the global sight of the whole parsing actions on the input sentence. Then we propose three kinds of ungreedy deterministic dependency parsing models for the modeling of parsing actions. The two original deterministic dependency parsers (Yamada's and Nivre's) are taken as baseline systems. Results show that ungreedy deterministic dependency parsers perform better than baseline systems while maintaining the same time complexity. The best ungreedy deterministic dependency parser improves much over baseline.

Deterministic Dependency Parsing Algorithms

Deterministic parsing algorithms regard parsing as a sequence of parsing actions that are taken step by step on the input sentence. Parsing actions construct dependency relations between words. In this paper, dependency tree is bare bone and dependency relations are just directed arcs. Because the set of parsing actions has limited number of elements, a classifier can be trained to classify parsing actions. During testing, parsing actions are determined by the trained classifier.

Ungreedy Deterministic Dependency Parsers

To overcome the greedy shortcoming of deterministic dependency parsers that choose the most probable parsing action at every step, we propose three kinds of ungreedy deterministic dependency parsing models: parsing action chain model, parsing action phrase model and n-phase model.

Parsing Action Chain Model (PACM)

The parsing process can be viewed as a Markov Chain. At every parsing step, there are several candidate parsing actions. The object is to find the most probable sequence of parsing actions by taking the Markov assumption. The classifier we use throughout this paper is SVM, which can output probabilities of parsing actions. Because this model chooses the most probable sequence, not the most probable parsing action at only one step, it avoids the greedy property of the original deterministic parsers.

Parsing Action Phrase Model (PAPM)

In the parsing action chain model, actions are competing at every parsing step. But for the parsing problem, it is reasonable that actions are competing for which phrase should be built. For dependency syntactic, one phrase consists of the head word and all its children. We propose parsing action phrase model, in which one parsing action phrase is comprised of a segment of consecutive parsing actions. It is segments of parsing actions that are competing for which next phrase should be built. The object is to find the most probable sequence of parsing action phrases.

N-Phase Model

If there are some circumstances in which the parsing actions are hard to determine, we can neglect these determinations in the first pass. When all other dependencies are constructed, clear contexts are gotten for those hard circumstances. Then a second pass is conducted to determine relations between remaining words. If there are n kinds of such hard circumstances, n+1 passes are needed. We call this as n-phase model. In Chinese, we find three kinds of hard circumstances, and the related experiments are carried out.

Experiments and Results

Experimental Setup

The data set for the experiments is taken from Penn Chinese Treebank (CHTB) version 5.0. Head rules (Sun and Jurafsky, 2004) are used to convert constituent structure to dependency structure.

The following metrics are used for evaluation:

Dependency accuracy (**DA**): The proportion of non-root words (excluding punctuations) that are assigned the correct head.

Root accuracy (**RA**): The proportion of root words that are correctly found.

Complete match (**CM**): The proportion of sentences whose dependency structures are completely correct.

Comparison of Ungreedy Deterministic Parser with Original Deterministic Parser, Generative Parser and Discriminative Parser

I implement the original deterministic dependency parsers proposed respectively by Yamada and Matsumoto (2003), Nivre and Scholz (2004). The generative parser used by me is dbparser, fulfilled by Daniel M. Bikel (2004). The discriminative dependency parser I use is MSTParser, fulfilled by Ryan Mcdonald (Mcdonald et al., 2005). The comparison of the ungreedy deterministic parser with all other parsers is presented in table 1. Although all ungreedy models perform better than original deterministic parsers, row ***ungreedy*** only lists the result of PAPM, which is the best among ungreedy models. From table 1, we can see that original deterministic parsers perform better than both generative and discriminative parsers, and the best ungreedy parser is much better than all other parsers.

	DA	RA	CM
ungreedy	**84.36**	**73.70**	**32.70**
Yamada	82.82	70.13	30.39
Nivre	82.69	68.19	29.82
dbparser	80.13	70.09	27.56
MSTParser	81.26	69.03	25.72

Table 1. Performances of all parsers.

Conclusion

This paper compares the deterministic dependency parsers with generative and discriminative parsers for Chinese dependency parsing. The results show that deterministic dependency parsers perform best. Based on the observation that deterministic parsers are greedy, we propose three kinds of ungreedy deterministic dependency parsers. The results show that ungreedy deterministic parsers perform much better than original deterministic parsers.

Ackowledgements

This work was supported by Hi-tech Research and Development Program of China under grant No. 2006AA01Z144, the Natural Sciences Foundation of China under grant No. 60673042, and the Natural Science Foundation of Beijing under grant No. 4052027, 4073043.

References

Daniel M. Bikel. 2004. *On the Parameter Space of Generative Lexicalized Statistical Parsing Models.* Ph.D. thesis, University of Pennsylvania.

Ryan McDonald, Koby Crammer, and Fernando Pereira. 2005. *Online large-margin training of dependency parsers.* In Proceedings of the 43rd AnnualMeeting of the Association for ComputationalLinguistics (ACL).

Joakim Nivre and Mario Scholz. 2004. *Deterministic dependency parsing of English text.* In Proceedings of the 20th COLING.

Kenji Sagae and Alon Lavie. 2005. *A classifier-based parser with linear run-time complexity.* In Proceedings of the 9th International Workshop on Parsing Technologies (IWPT).

Honglin Sun and Daniel Jurafsky. 2004. *Shallow semantic parsing of Chinese.* In Proceedings of the HLT/NAACL'04.

Mengqiu Wang, Kenhi Sagae, and Teruko Mitamura. 2006. *A fast, accurate deterministic parser for Chinese.* In Proceedings of the 44th AnnualMeeting of the Association for ComputationalLinguistics (ACL).

Hiroyasu Yamada and Yuji Matsumoto. 2003. *Statistical dependency analysis with support vector machines.* In Proceedings of the 8th InternationalWorkshop on Parsing Technologies (IWPT).