Investigating the Impact of Syntactic Features for Semantic Dependency Parsing

Maximilian Pfundstein

January 6, 2020

Linköpings University

Table of contents

Introduction

Introduction | Research Question

- Syntatic and Semantic Dependency Parsing (SDP)
- Acyclic Graphs
- Research Question: To what extend does predicted syntax help for semantic dependency parsing?

- Theoretical baseline model by Dozat and Manning [Zeman et al., 2017]
- Implementation by Daniel Roxbo [Roxbo, 2019]
- Extension to learn syntactic graphs

Semantic Only	Gold Syntax	Predicted Syntax
93.6%	97.1%	92.8%

Table 1: Semantic F_1 scores.

- Multitask Learning
- Presumption by [Kurtz et al., 2019] shows that there might be an information overlap
- Other research shows, that multitask learning might help [Peng et al., 2017]

Overview

Overview | Objectives

- Allow to import the data from the `.conllu` format
- Increase the performance during the pre-processing by using caching to save computation time
- Implement multitask learning to simultaneously learn syntactic and semantic graphs and investigate the results

Objectives

Objectives | Import `.conllu`

- Separate files for syntax (`.cpn`) and semantic (`.sdp`)
- Combined file format CoNLL-U from [Buchholz and Marsi, 2006]

Objectives | Performance Improvement

- Preprocessing is conducted every time the source code is being invoked
- Main functions:
  - `parse_conllu_labels`
  - `parse_conllu_sentences`
  - `parse_conllu_targets`

Objectives | Multitask Learning

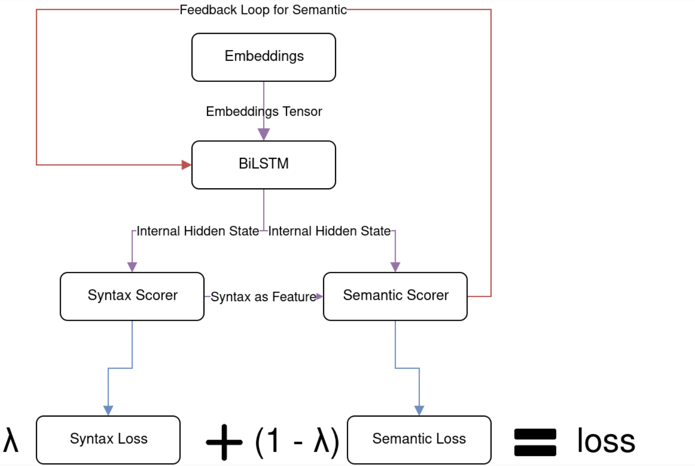
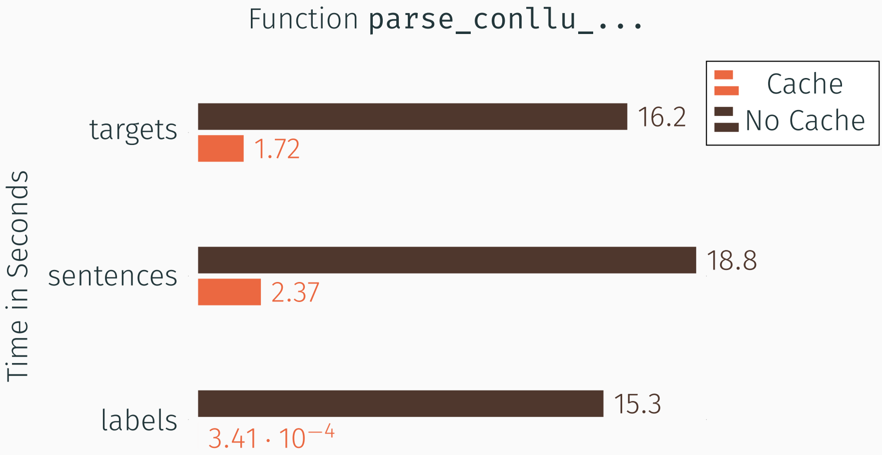


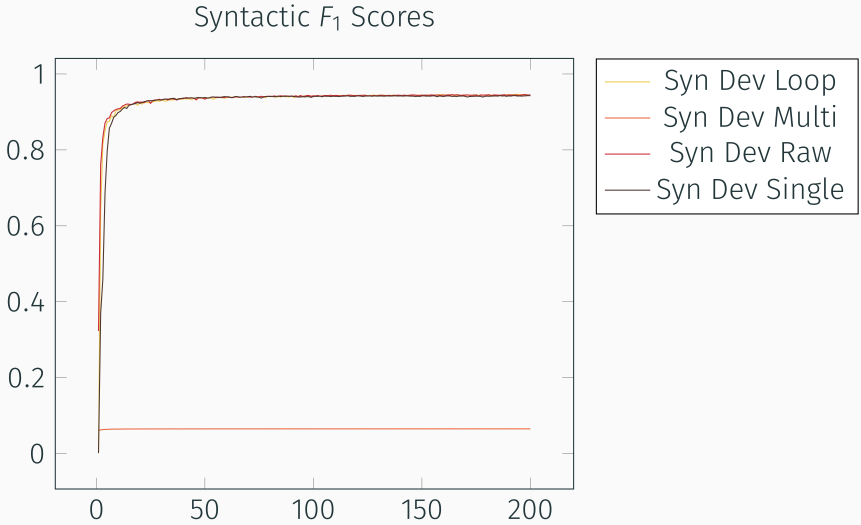
Figure 1: Flowchart, with red representing the feedback loop.

Results

Results | Performance Improvement

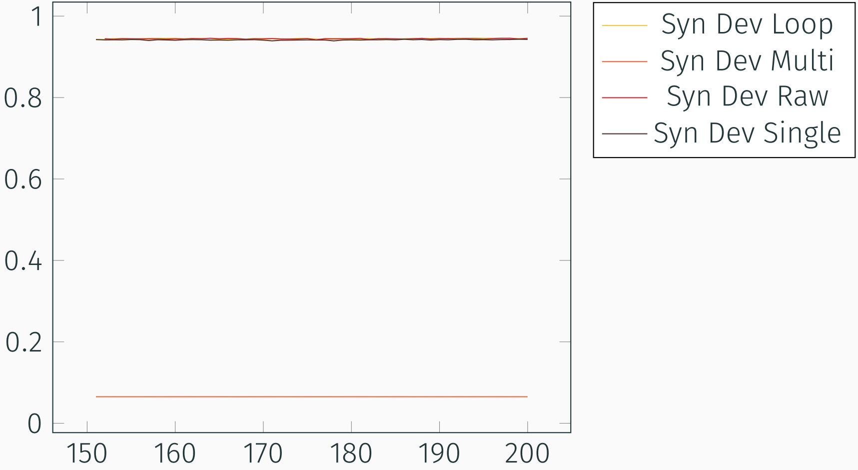


Results | Multitask Learning

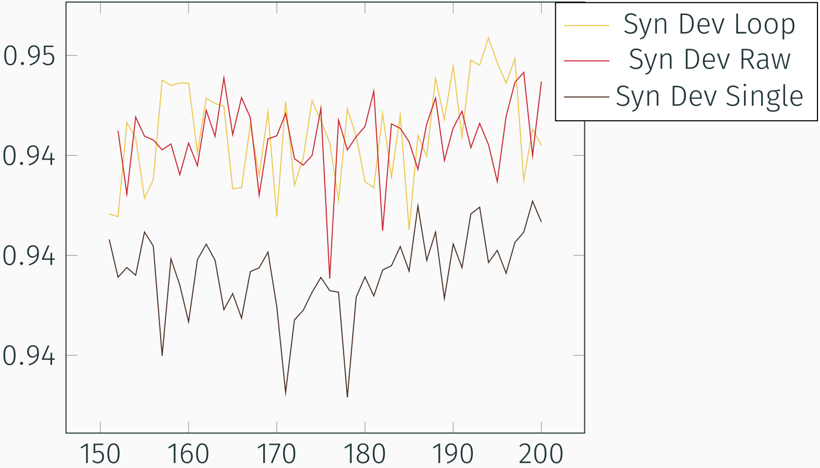


Multitask Learning

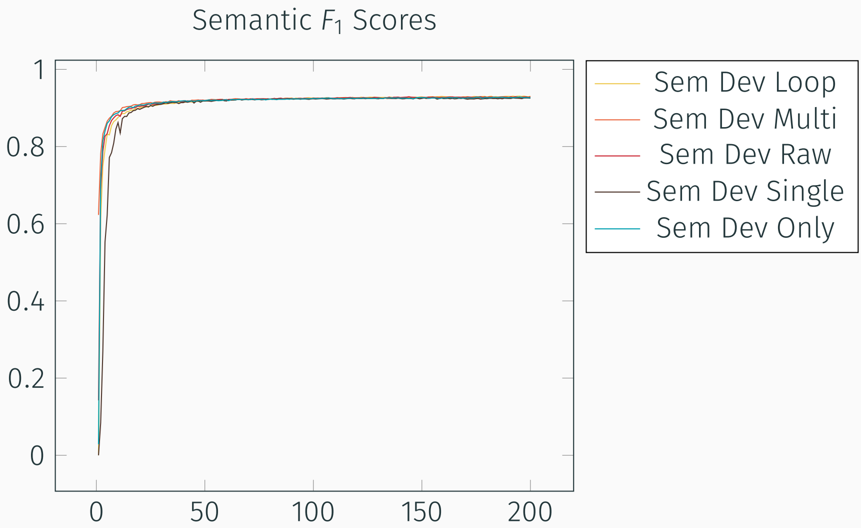
Syntactic F_1 Scores (last 50 epochs)



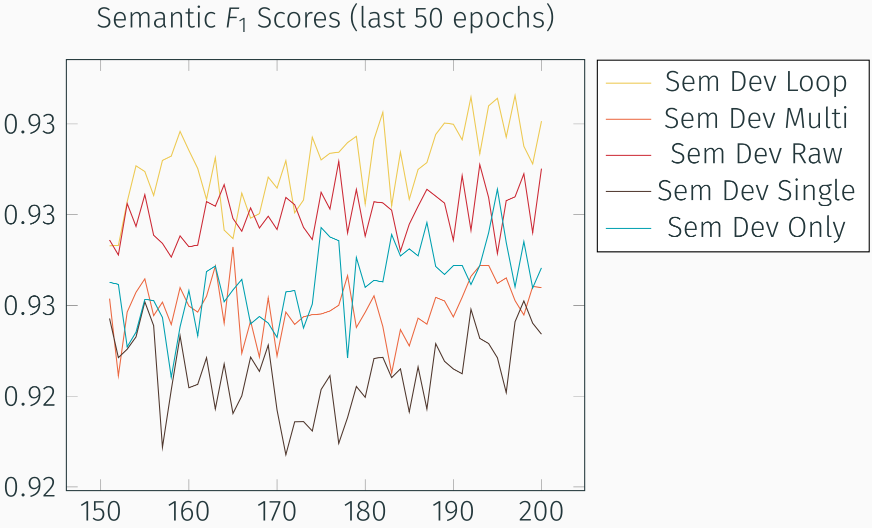
Syntactic F_1 Scores (last 50 epochs without Syn Dev Multi)



Results | Multitask Learning



Results | Multitask Learning



Setting	**Syn Dev**	**Sem Dev**	**Syn Test**	**Sem Test**
Sem Only	0	92.39%	0	**87.50%**
Multitask Raw	-	-	90.92%	**87.80%**
Multitask Single	94.64%	92.6%	90.81%	87.41%
Multitask Multi	-	-	7.39%	87.71%
Multitask Loop	94.29%	92.51%	90.96%	**87.87%**

Table 2: F_1 scores for the different setups.

Outlook

Outlook

- Presumption by [Kurtz et al., 2019] confirmed, at least for this data set
- Bidirectional Encoder Representations from Transformers (BERT)
- Combine BERT and multitask learning

References

- Buchholz, S. and Marsi, E. (2006).
  Conll-x shared task on multilingual dependency parsing.
  In *Proceedings of the Tenth Conference on Computational Natural Language Learning*, CoNLL-X '06, page 149–164, USA. Association for Computational Linguistics.
- Kurtz, R., Roxbo, D., and Kuhlmann, M. (2019).
  Improving semantic dependency parsing with syntactic features.
  In *Proceedings of the First NLPL Workshop on Deep Learning for Natural Language Processing*, pages 12–21, Turku, Finland. Linköping University Electronic Press.

Peng, H., Thomson, S., and Smith, N. A. (2017).
Deep multitask learning for semantic dependency parsing.
In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2037–2048, Vancouver, Canada. Association for Computational Linguistics.

Roxbo, D. (2019).
A detailed analysis of semantic dependency parsing with deep neural networks.
Master’s thesis, Linköping University, Human-Centered systems.

Zeman, D., Popel, M., Straka, M., Hajič, J., Nivre, J., Ginter, F., Luotolahti, J., Pyysalo, S., Petrov, S., Potthast, M., Tyers, F., Badmaeva, E., Gokirmak, M., Nedoluzhko, A., Cinková, S., Hajič jr., J., Hlaváčová, J., Kettnerová, V., Urešová, Z., Kanerva, J., Ojala, S., Missilä, A., Manning, C. D., Schuster, S., Reddy, S., Taji, D., Habash, N., Leung, H., de Marneffe, M.-C., Sanguinetti, M., Simi, M., Kanayama, H., de Paiva, V., Droganova, K., Martínez Alonso, H., Çöltekin, Ç., Sulubacak, U., Uszkoreit, H., Macketanz, V., Burchardt, A., Harris, K., Marheinecke, K., Rehm, G., Kayadelen, T., Attia, M., Elkahky, A., Yu, Z., Pitler, E., Lertpradit, S., Mandl, M., Kirchner, J., Alcalde, H. F., Strnadová, J., Banerjee, E., Manurung, R.,

Stella, A., Shimada, A., Kwak, S., Mendonça, G., Lando, T., Nitisaroj, R., and Li, J. (2017).
CoNLL 2017 shared task: Multilingual parsing from raw text to universal dependencies.
In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 1–19, Vancouver, Canada. Association for Computational Linguistics.