CoNLL 2018

The SIGNLL Conference on Computational Natural Language Learning

Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies

October 31 – November 1, 2018
Brussels, Belgium

Sponsors:

Google, Inc.

DFKI Berlin

CRACKER project

text & form

ÚFAL

Charles University

LINDAT/CLARIN

Czech Science Foundation
project No. 15-10472S

©2018 The Association for Computational Linguistics

ISBN 978-1-948087-82-7

Introduction

This volume contains papers describing systems submitted to the *CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies,* and two overview papers: one summarizing the main task, its features, evaluation methodology for the main and additional metrics, and some interesting observations about the submitted systems and the task as a whole; the other overview paper discusses a complementary task focusing on *Extrinsic Parser Evaluation (EPE)*.

This Shared Task (`http://universaldependencies.org/conll18/`) is an extension of the CoNLL 2017 Shared Task, with certain important differences. Like in the previous year, the data come from the Universal Dependencies project (`http://universaldependencies.org`), which provides annotated treebanks for a large number of languages using the same annotation scheme. The number of treebanks in the task (82) is similar to the previous year (81), but some treebanks from the 2017 task were not included in the present task, and some new treebanks were added instead. The datasets are samples from 57 different languages (all languages from the previous year, and eight new languages). In comparison to 2017, there were more low-resource languages with extremely little training data, calling for cross-lingual transfer techniques. Unlike 2017, none of the low-resource languages were "surprise languages".

Participants had to process all the test sets. The TIRA platform has been used for evaluation, as was the case already for the CoNLL 2015, 2016 and 2017 Shared Tasks, meaning that participants had to provide their code on a designated virtual machine to be run by the organizers to produce official results. However, test data have been published after the official evaluation period, and participants could run their systems at home to produce additional results they were allowed to include in the system description papers.

The systems were ranked according to three main evaluation metrics – LAS (Labeled Attachment Score), MLAS (Morphology-aware Labeled Attachment Score), and BLEX (BiLEXical Dependency Score). Like last year, participating systems minimally had to find labeled syntactic dependencies between words. In addition, this year's task featured new metrics that also scored a system's capacity to predict a morphological analysis of each word, including a part-of-speech tag, morphological features, and a lemma. Regardless of metric, the assumption was that the input should be raw text, with no gold-standard word or sentence segmentation, and no gold-standard morphological annotation. However, for teams who wanted to concentrate on one or more subtasks, segmentation and morphology predicted by a baseline system was made available.

The complementary EPE task seeks to provide better estimates of the relative utility of different parsers for a variety of downstream applications that depend centrally on the analysis of grammatical structure, viz. biomedical event extraction, negation resolution, and fine-grained opinion analysis. EPE 2018 was organized as an optional add-on exercise to the core task: the submitted systems were applied to extra texts, about 1.1 million tokens of English. It is interesting to see to what degree different intrinsic evaluation metrics from the core task correlate with end-to-end EPE results, and comparison to earlier EPE campaigns—with other types of dependency representations and additional sources of training data—further helps to put the core task into perspective.

A total of 25 systems ran successfully and have been ranked (`http://universaldependencies.org/conll18/results.html`); 17 teams submitted parser outputs for the EPE test set. While there are clear overall winners in each of the evaluation metrics, we would like to thank all participants for working hard on their submissions and adapting their systems not only to the datasets available, but also to the evaluation platform. We would like to thank all of them for their effort, since it is the participants who are the core of any shared task's success.

We would like to thank the CoNLL organizers for their support and the reviewers for helping to improve the submitted system papers. Special thanks go to Martin Potthast of the TIRA platform for handling such a large number of systems, running often for several hours each, and for being very responsive and helpful to us and all system participants, round the clock during the evaluation phase and beyond. We also thank to the 200+ people working on the Universal Dependencies project during the past four years, without whom there would be no data.

Daniel Zeman, Jan Hajič, Martin Popel, Milan Straka, Joakim Nivre, Filip Ginter and Slav Petrov, the organizers of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies
Stephan Oepen, main organizer of the EPE 2018 task

Prague, September 2018

Chairs:

Daniel Zeman, Charles University, Prague
Jan Hajič, Charles University, Prague

Support Team:

Joakim Nivre, Uppsala University
Filip Ginter, University of Turku
Slav Petrov, Google
Milan Straka, Charles University, Prague
Martin Popel, Charles University, Prague

TIRA support:

Martin Potthast, Leipzig University

Data Providers:

We are immensely grateful to the maintainers and contributors of the Universal Dependencies treebanks used in this shared task.

Extrinsic Evaluation Liaisons:

Jari Björne, University of Turku
Murhaf Fares, University of Oslo
Stephan Oepen, University of Oslo

Program Committee:

Miguel Ballesteros, IBM Research AI
James Barry, ADAPT Centre DCU
Claudia Borg, University of Malta
Wanxiang Che, Harbin Institute of Technology
Danlu Chen, Fudan University
Miryam de Lhoneux, Uppsala University
Agnieszka Faleńska, Universität Stuttgart
Marcos Garcia, Universidade da Corunha
Filip Ginter, University of Turku
Can Gümeli, Koç University
Shexia He, Shanghai Jiao Tong University
Daniel Hershcovich, Hebrew University
Jenna Kanerva, University of Turku
Hrant Khachatrian, YerevaNN
Ömer Kırnap, Koç University
Zuchao Li, Shanghai Jiao Tong University
KyungTae Lim, LATTICE-CNRS
Mengxiao Lin, Megvii Technology
Olivier Mesnard, CEA LIST
Dang Nguyen, University of Information Technology, VNU-HCM
Dat Quoc Nguyen, The University of Melbourne
Joakim Nivre, Uppsala University
Atul Kr. Ojha, SSIS, JNU
Cheoneum Park, Kangwon National University
Martin Popel, Charles University
Peng Qi, Stanford University
Berkay Önder, Koç University
Şaziye Betül Özateş, Bogazici University
Rudolf Rosa, Charles University
Piotr Rybak, ICS PAS
Djamé Seddah, Université Paris Sorbonne
Miikka Silfverberg, University of Colorado
Francis Tyers, Indiana University
Hui Wan, IBM Research AI
Yingting Wu, Shanghai Jiao Tong University
Xiang Yu, IMS, University of Stuttgart
Zhuosheng Zhang, Shanghai Jiao Tong University
Pengfei Zhu, Tianjin Univerity

Table of Contents

CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies
Daniel Zeman, Jan Hajič, Martin Popel, Martin Potthast, Milan Straka, Filip Ginter, Joakim Nivre and Slav Petrov 1

The 2018 Shared Task on Extrinsic Parser Evaluation: On the Downstream Utility of English Universal Dependency Parsers
Murhaf Fares, Stephan Oepen, Lilja Øvrelid, Jari Björne and Richard Johansson 22

CEA LIST: Processing Low-Resource Languages for CoNLL 2018
Elie Duthoo and Olivier Mesnard 34

Semi-Supervised Neural System for Tagging, Parsing and Lematization
Piotr Rybak and Alina Wróblewska 45

Towards Better UD Parsing: Deep Contextualized Word Embeddings, Ensemble, and Treebank Concatenation
Wanxiang Che, Yijia Liu, Yuxuan Wang, Bo Zheng and Ting Liu 55

Joint Learning of POS and Dependencies for Multilingual Universal Dependency Parsing
Zuchao Li, Shexia He, Zhuosheng Zhang and Hai Zhao 65

Multilingual Universal Dependency Parsing from Raw Text with Low-Resource Language Enhancement
Yingting Wu, Hai Zhao and Jia-Jun Tong 74

An Improved Neural Network Model for Joint POS Tagging and Dependency Parsing
Dat Quoc Nguyen and Karin Verspoor 81

IBM Research at the CoNLL 2018 Shared Task on Multilingual Parsing
Hui Wan, Tahira Naseem, Young-Suk Lee, Vittorio Castelli and Miguel Ballesteros 92

Universal Dependency Parsing with a General Transition-Based DAG Parser
Daniel Hershcovich, Omri Abend and Ari Rappoport 103

82 Treebanks, 34 Models: Universal Dependency Parsing with Multi-Treebank Models
Aaron Smith, Bernd Bohnet, Miryam de Lhoneux, Joakim Nivre, Yan Shao and Sara Stymne . 113

Tree-Stack LSTM in Transition Based Dependency Parsing
Ömer Kırnap, Erenay Dayanık and Deniz Yuret 124

Turku Neural Parser Pipeline: An End-to-End System for the CoNLL 2018 Shared Task
Jenna Kanerva, Filip Ginter, Niko Miekka, Akseli Leino and Tapio Salakoski 133

SEx BiST: A Multi-Source Trainable Parser with Deep Contextualized Lexical Representations
KyungTae Lim, Cheoneum Park, Changki Lee and Thierry Poibeau 143

The SLT-Interactions Parsing System at the CoNLL 2018 Shared Task
Riyaz A. Bhat, Irshad Bhat and Srinivas Bangalore 153

Universal Dependency Parsing from Scratch
Peng Qi, Timothy Dozat, Yuhao Zhang and Christopher D. Manning 160

NLP-Cube: End-to-End Raw Text Processing With Neural Networks
Tiberiu Boroș, Stefan Daniel Dumitrescu and Ruxandra Burtica 171

Towards JointUD: Part-of-speech Tagging and Lemmatization using Recurrent Neural Networks
Gor Arakelyan, Karen Hambardzumyan and Hrant Khachatrian 180

CUNI x-ling: Parsing Under-Resourced Languages in CoNLL 2018 UD Shared Task
Rudolf Rosa and David Mareček 187

UDPipe 2.0 Prototype at CoNLL 2018 UD Shared Task
Milan Straka 197

Universal Morpho-Syntactic Parsing and the Contribution of Lexica: Analyzing the ONLP Lab Submission to the CoNLL 2018 Shared Task
Amit Seker, Amir More and Reut Tsarfaty 208

SParse: Koç University Graph-Based Parsing System for the CoNLL 2018 Shared Task
Berkay Önder, Can Gümeli and Deniz Yuret 216

ELMoLex: Connecting ELMo and Lexicon Features for Dependency Parsing
Ganesh Jawahar, Benjamin Muller, Amal Fethi, Louis Martin, Eric Villemonte de la Clergerie, Benoît Sagot and Djamé Seddah 223

A Morphology-Based Representation Model for LSTM-Based Dependency Parsing of Agglutinative Languages
Şaziye Betül Özateş, Arzucan Özgür, Tunga Gungor and Balkız Öztürk 238

AntNLP at CoNLL 2018 Shared Task: A Graph-Based Parser for Universal Dependency Parsing
Tao Ji, Yufang Liu, Yijun Wang, Yuanbin Wu and Man Lan 248

A Simple yet Effective Joint Training Method for Cross-Lingual Universal Dependency Parsing
Danlu Chen, Mengxiao Lin, Zhifeng Hu and Xipeng Qiu 256