IWPT 2021

The 17th International Conference on Parsing Technologies

Proceedings of the Conference (Including the IWPT 2021 Shared Task)

August 6, 2021
Bangkok, Thailand (online)

©2021 The Association for Computational Linguistics
and The Asian Federation of Natural Language Processing

Order copies of this and other ACL-IJCNLP proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-954085-80-0

Preface

Welcome to the 17th International Conference on Parsing Technologies (IWPT 2021), which this year (for only the second time since 2007) is co-located with the Annual Meeting of the Association for Computational Linguistics and of the Asian Federation of Natural Language Processing (ACL-IJCNLP). The IWPT meeting series, hosted by the ACL Special Interest Group in Natural Language Parsing (SIGPARSE), has been held biennualy since its inaugual meeting in 1989 in Pittsburgh, PA (USA).

For 2021, the SIGPARSE steering group decided to continue an experiment started last year, co-location with the main ACL meeting in the form of a reduced one-day IWPT programme. The main motivation for this move was to reduce fragmentation (and travel) and to increase IWPT visibility in the broader ACL community. We believe that both these goals have been attained in 2020 and hope to see this development continue this year. At the same time, IWPT has launched its own series of parsing shared tasks since 2020, which strengthens the experimental and applied perspective on parsing technologies in the conference programme.

The IWPT 2021 shared task focuses on the parsing of Enhanced Universal Dependencies (EUD) over 17 languages, continuing on from a successful EUD parsing shared task in 2020. This is the second time that graph-based representations of syntactic structures are evaluated on such a large scale, and we believe it will pave the way for research on richer models and representations. The shared task attracted system submissions from nine teams from around the world and, thus, establishes a highly relevant point of comparison for this line of syntactic analysis. We are very grateful to everyone who contributed to this shared task, starting with the data providers who worked hard to meet our deadline. Thanks to the participant teams who worked tirelessly in a short time period to provide such a set of great and interesting systems!

Owing to the ongoing pandemic, the meeting will regrettably once again be held entirely virtual, where for IWPT we have adopted a mostly-asynchronous format: Accepted papers (of four different types, long, short, shared task, and findings) will be presented through pre-recorded talks, which become available on-line for individual viewing before the actual conference day. On the day of the conference, August 6, there will be a four-hour live session, scheduled so that the timing should be convenient (all things considered) for participants around the world: 13:00–17:00 UTC+0, which translates, for example, into a starting time at 6:00 in the morning at the US West Coast and wrapping up at 1:00 in the morning in Melbourne, Australia. The live sessions will be devoted exclusively to questions and answers, organized into six thematic sessions. Authors of papers associated with each session will be available to answer questions and discuss their work (possibly also among themselves).

There has been (and to some degree still is) much uncertainty about the format of ACL-IJCNLP and IWPT this year, and in a sense we were positively surprised to receive a number of submissions comparable to recent IWPT instances. Out of 24 regular paper submissions, the programme committee accepted 13 for presentation at the conference. The IWPT 2021 programme is complemented by one invited talk, by Emily Pitler of Google Research (to whom we are immensely grateful for honoring her commitment despite the mostly-asynchronous, virtual format), by four papers adopted from the Findings of ACL-IJCNLP 2021, and by an overview paper and nine system descriptions from the IWPT 2021 shared task. We further gratefully acknowledge the work, flexibility, and collegiality of authors and reviewers, as well as of the ACL-IJCNLP workshop and publication chairs, who had to shepherd our community through a difficult logistics process.

Davis, Groningen, Oslo, Paris, Prague, and Tel Aviv

Gosse Bouma, Stephan Oepen, Kenji Sagae,
Djamé Seddah, Reut Tsarfaty, and Dan Zeman

Organizers:

Stephan Oepen, University of Oslo (General Co-Chair)
Kenji Sagae, University of California at Davis (General Co-Chair)
Reut Tsarfaty, Open University of Israel (General Co-Chair)
Gosse Bouma, University of Groningen (Shared Task Co-Chair)
Djamé Seddah, University Paris-Sorbonne (Shared Task Co-Chair)
Dan Zeman, Charles University in Prague (Shared Task Co-Chair)

Program Committee:

Omri Abend
Željko Agić
Mark Anderson
Miguel Ballesteros
James Barry
Gosse Bouma
Marie Candito
Xavier Carreras
John Carroll
Maximin Coavoux
Éric de la Clergerie
Miryam de Lhoneux
Agnieszka Falenska
Jennifer Foster
Annemarie Friedrich
Stefan Grünewald
Carlos Gómez-Rodríguez
Zixia Jia
Yong Jiang
Lifeng Jin
Nikita Kitaev
Mateusz Klimaszewski
Marco Kuhlmann
Jonathan K. Kummerfeld
Sandra Kübler
Joseph Le Roux
Yusuke Miyao
Mark-Jan Nederhof
Joakim Nivre
Stephan Oepen
Barbara Plank
Roi Reichart
Corentin Ribeyre
Kenji Sagae
Giorgio Satta
Natalie Schluter
Djamé Seddah
Tianze Shi
Reut Tsarfaty

Gertjan van Noord
David Vilares
Daniel Zeman
Yi Zhang
Yue Zhang
Lilja Øvrelid

Invited Speaker:

Emily Pitler, Google Research

Table of Contents

Generic Oracles for Structured Prediction
Christoph Teichmann and Antoine Venant 1

Proof Net Structure for Neural Lambek Categorial Parsing
Aditya Bhargava and Gerald Penn 13

The Reading Machine: A Versatile Framework for Studying Incremental Parsing Strategies
Franck Dary and Alexis Nasr 26

Semi-Automatic Construction of Text-to-SQL Data for Domain Transfer
Tianyi Li, Sujian Li and Mark Steedman 38

Levi Graph AMR Parser using Heterogeneous Attention
Han He and Jinho D. Choi 50

Translate, then Parse! A Strong Baseline for Cross-Lingual AMR Parsing
Sarah Uhrig, Yoalli Garcia, Juri Opitz and Anette Frank 58

Great Service! Fine-grained Parsing of Implicit Arguments
Ruixiang Cui and Daniel Hershcovich 65

A Falta de Pan, Buenas Son Tortas: The Efficacy of Predicted UPOS Tags for Low Resource UD Parsing
Mark Anderson, Mathieu Dehouck and Carlos Gómez-Rodríguez 78

Multilingual Dependency Parsing for Low-Resource African Languages: Case Studies on Bambara, Wolof, and Yoruba
Cheikh M. Bamba Dione 84

Bidirectional Domain Adaptation Using Weighted Multi-Task Learning
Daniel Dakota, Zeeshan Ali Sayyed and Sandra Kübler 93

Strength in Numbers: Averaging and Clustering Effects in Mixture of Experts for Graph-Based Dependency Parsing
xudong zhang, Joseph Le Roux and Thierry Charnois 106

A Modest Pareto Optimisation Analysis of Dependency Parsers in 2021
Mark Anderson and Carlos Gómez-Rodríguez 119

Applying Occam's Razor to Transformer-Based Dependency Parsing: What Works, What Doesn't, and What is Really Necessary
Stefan Grünewald, Annemarie Friedrich and Jonas Kuhn 131

Incorporating Compositionality and Morphology into End-to-End Models
Emily Pitler 145

From Raw Text to Enhanced Universal Dependencies: The Parsing Shared Task at IWPT 2021
Gosse Bouma, Djamé Seddah and Daniel Zeman 146

COMBO: A New Module for EUD Parsing
Mateusz Klimaszewski and Alina Wróblewska 158

Splitting EUD Graphs into Trees: A Quick and Clatty Approach
Mark Anderson and Carlos Gómez-Rodríguez 167

Graph Rewriting for Enhanced Universal Dependencies
Bruno Guillaume and Guy Perrier 175

Biaffine Dependency and Semantic Graph Parsing for EnhancedUniversal Dependencies
Giuseppe Attardi, Daniele Sartiano and Maria Simi 184

Enhanced Universal Dependency Parsing with Automated Concatenation of Embeddings
Xinyu Wang, Zixia Jia, Yong Jiang and Kewei Tu 189

RobertNLP at the IWPT 2021 Shared Task: Simple Enhanced UD Parsing for 17 Languages
Stefan Grünewald, Frederik Tobias Oertel and Annemarie Friedrich 196

The DCU-EPFL Enhanced Dependency Parser at the IWPT 2021 Shared Task
James Barry, Alireza Mohammadshahi, Joachim Wagner, Jennifer Foster and James Henderson204

TGIF: Tree-Graph Integrated-Format Parser for Enhanced UD with Two-Stage Generic- to Individual-Language Finetuning
Tianze Shi and Lillian Lee 213

End-to-end mBERT based Seq2seq Enhanced Dependency Parser with Linguistic Typology knowledge
Chinmay Choudhary and Colm O'riordan 225

Conference Program

Friday, August 6, 2021 (All Times in UTC+0)

13:00–13:30 Session 1: Regular Papers Q&A

Generic Oracles for Structured Prediction
Christoph Teichmann and Antoine Venant

Proof Net Structure for Neural Lambek Categorial Parsing
Aditya Bhargava and Gerald Penn

Findings: Neural Combinatory Constituency Parsing
Zhousi Chen, Longtu Zhang, Aizhan Imankulova, Mamoru Komachi

Findings: Rule Augmented Unsupervised Constituency Parsing
Atul Sahay, Anshul Nasery, Ayush Maheshwari, Ganesh Ramakrishnan, Rishabh Iyer

13:30–14:05 Session 2: Regular Papers Q&A

The Reading Machine: A Versatile Framework for Studying Incremental Parsing Strategies
Franck Dary and Alexis Nasr

Semi-Automatic Construction of Text-to-SQL Data for Domain Transfer
Tianyi Li, Sujian Li and Mark Steedman

Levi Graph AMR Parser using Heterogeneous Attention
Han He and Jinho D. Choi

Translate, then Parse! A Strong Baseline for Cross-Lingual AMR Parsing
Sarah Uhrig, Yoalli Garcia, Juri Opitz and Anette Frank

Great Service! Fine-grained Parsing of Implicit Arguments
Ruixiang Cui and Daniel Hershcovich

Friday, August 6, 2021 (All Times in UTC+0) (continued)

14:05–14:35 Session 3: Regular Papers Q&A

A Falta de Pan, Buenas Son Tortas: The Efficacy of Predicted UPOS Tags for Low Resource UD Parsing
Mark Anderson, Mathieu Dehouck and Carlos Gómez-Rodríguez

Multilingual Dependency Parsing for Low-Resource African Languages: Case Studies on Bambara, Wolof, and Yoruba
Cheikh M. Bamba Dione

Findinigs: Climbing the Tower of Treebanks: Improving Low-Resource Dependency Parsing via Hierarchical Source Selection
Goran Glavaš, Ivan Vulić

Findings: Annotations Matter: Leveraging Multi-task Learning to Parse UD and SUD
Zeeshan Ali Sayyed, Daniel Dakota

14:35–15:05 Session 5: Regular Papers Q&A

Bidirectional Domain Adaptation Using Weighted Multi-Task Learning
Daniel Dakota, Zeeshan Ali Sayyed and Sandra Kübler

Strength in Numbers: Averaging and Clustering Effects in Mixture of Experts for Graph-Based Dependency Parsing
xudong zhang, Joseph Le Roux and Thierry Charnois

A Modest Pareto Optimisation Analysis of Dependency Parsers in 2021
Mark Anderson and Carlos Gómez-Rodríguez

Applying Occam's Razor to Transformer-Based Dependency Parsing: What Works, What Doesn't, and What is Really Necessary
Stefan Grünewald, Annemarie Friedrich and Jonas Kuhn

Friday, August 6, 2021 (All Times in UTC+0) (continued)

15:05–15:15 Break

15:15–16:00 Session 4: Invited Talk

Incorporating Compositionality and Morphology into End-to-End Models
Emily Pitler

16:00–16:45 Session 6: Shared Task Q&A

From Raw Text to Enhanced Universal Dependencies: The Parsing Shared Task at IWPT 2021
Gosse Bouma, Djamé Seddah and Daniel Zeman

COMBO: A New Module for EUD Parsing
Mateusz Klimaszewski and Alina Wróblewska

Splitting EUD Graphs into Trees: A Quick and Clatty Approach
Mark Anderson and Carlos Gómez-Rodríguez

Graph Rewriting for Enhanced Universal Dependencies
Bruno Guillaume and Guy Perrier

Biaffine Dependency and Semantic Graph Parsing for EnhancedUniversal Dependencies
Giuseppe Attardi, Daniele Sartiano and Maria Simi

Enhanced Universal Dependency Parsing with Automated Concatenation of Embeddings
Xinyu Wang, Zixia Jia, Yong Jiang and Kewei Tu

RobertNLP at the IWPT 2021 Shared Task: Simple Enhanced UD Parsing for 17 Languages
Stefan Grünewald, Frederik Tobias Oertel and Annemarie Friedrich

The DCU-EPFL Enhanced Dependency Parser at the IWPT 2021 Shared Task
James Barry, Alireza Mohammadshahi, Joachim Wagner, Jennifer Foster and James Henderson

TGIF: Tree-Graph Integrated-Format Parser for Enhanced UD with Two-Stage Generic- to Individual-Language Finetuning
Tianze Shi and Lillian Lee

16:45–17:00 Session 7: SIGPARSE Business Meeting

End-to-end mBERT based Seq2seq Enhanced Dependency Parser with Linguistic Typology knowledge
Chinmay Choudhary and Colm O'riordan