

EMNLP 2020

Structured Prediction for NLP

Proceedings of the Fourth Workshop

November 20, 2020

©2020 The Association for Computational Linguistics

Order copies of this and other ACL 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 978-1-952148-83-5

Introduction

Welcome to the Fourth Workshop on Structured Prediction for NLP!

Structured prediction has a strong tradition within the natural language processing (NLP) community, owing to the discrete, compositional nature of words and sentences, which leads to natural combinatorial representations such as trees, sequences, segments, or alignments, among others. It is no surprise that structured output models have been successful and popular in NLP applications since their inception. Many other NLP tasks, including, but not limited to: semantic parsing, slot filling, machine translation, or information extraction, are commonly modeled as structured problems, and accounting for said structure has often lead to performance gain.

Of late, continuous representation learning via neural networks has been a significant complementary direction, leading to improvements in unsupervised and semi-supervised pre-training, transfer learning, domain adaptation, etc. Using word embeddings as features for structured models such as part-of-speech taggers count among the very first uses of continuous embeddings in NLP, and the symbiosis between the two approaches is an exciting research direction today.

This year we received 26 submissions and, after double-blind peer review, 16 were accepted (4 of which are non-archival papers) for presentation in this edition of the workshop, all exploring this interplay between structure and neural data representations, from different, important points of view. The program includes work on structure-informed representation learning, energy-based learning, and structured fine-tuning of language models. Our program also includes six invited presentations from influential researchers.

Our warmest thanks go to the program committee – for their time and effort providing valuable feedback, to all submitting authors – for their thought-provoking work, and to the invited speakers – for doing us the honor of joining our program. We are looking forward to seeing you online!

Priyanka Agrawal
Zornitsa Kozareva
Julia Kreutzer
Gerasimos Lampouras
André Martins
Sujith Ravi
Andreas Vlachos

Organizers:

Priyanka Agrawal, Booking.com, Netherlands
Zornitsa Kozareva, Google, USA
Julia Kreutzer, Google, Canada
Gerasimos Lampouras, Huawei, UK
André F. T. Martins, Unbabel and Instituto de Telecomunicações, Portugal
Sujith Ravi, Amazon, USA
Andreas Vlachos, University of Cambridge, UK

Program Committee:

Sweta Agrawal, University of Maryland, USA
Naveen Arivazhagan, Google Research, USA
Yoav Artzi, Cornell University, USA
Wilker Aziz, University of Amsterdam, Netherlands
Colin Cherry, Google Research, Canada
Gonçalo Correia, Instituto de Telecomunicacoes, Portugal
George Foster, Google Research, Canada
Kevin Gimpel, TTI Chicago, USA
Parag Jain, University of Edinburgh, UK
Arzoo Katiyar, Cornell University, USA
Yoon Kim, MIT-IBM Watson AI Lab, USA
Parisa Kordjamshidi, Tulane University, USA
Chunchuan Lyu, University of Edinburgh, UK
Pranava Madhyastha, Imperial College London, UK
Zita Marinho, Sacoor Brothers, Portugal
Musie Meressa, Sapienza University of Rome, Italy
Sabrina J Mielke, Johns Hopkins University, USA
Toan Q Nguyen, University of Notre Dame, USA
Vlad Niculae, University of Amsterdam, Netherlands
Marek Rei, University of Cambridge, UK
Hiko Schamoni, Heidelberg University, Germany
Tianze Shi, Cornell University, USA
Vivek Srikumar, University of Utah, USA
Sean J Welleck, New York University, USA

Invited Speakers:

Isabelle Augenstein, University of Copenhagen
Jonathan Berant, Tel-Aviv University
Mark Johnson, Macquarie University
Alexander Rush, Cornell Tech
Sunita Sarawagi, IIT Bombay
Ivan Titov, University of Edinburgh

Table of Contents

<i>Syntax-driven Iterative Expansion Language Models for Controllable Text Generation</i> Noe Casas, José A. R. Fonollosa and Marta R. Costa-jussà	1
<i>CopyNext: Explicit Span Copying and Alignment in Sequence to Sequence Models</i> Abhinav Singh, Patrick Xia, Guanghui Qin, Mahsa Yarmohammadi and Benjamin Van Durme .	11
<i>Generating Synthetic Data for Task-Oriented Semantic Parsing with Hierarchical Representations</i> Ke Tran and Ming Tan	17
<i>Structured Prediction for Joint Class Cardinality and Entity Property Inference in Model-Complete Text Comprehension</i> Hendrik ter Horst and Philipp Cimiano	22
<i>Energy-based Neural Modelling for Large-Scale Multiple Domain Dialogue State Tracking</i> Anh Duong Trinh, Robert J. Ross and John D. Kelleher	33
<i>End-to-End Extraction of Structured Information from Business Documents with Pointer-Generator Networks</i> Clément Sage, Alex Aussem, Véronique Eglin, Haytham Elghazel and Jérémy Espinas	43
<i>Layer-wise Guided Training for BERT: Learning Incrementally Refined Document Representations</i> Nikolaos Manginas, Ilias Chalkidis and Prodromos Malakasiotis	53
<i>Improving Joint Training of Inference Networks and Structured Prediction Energy Networks</i> Lifu Tu, Richard Yuanzhe Pang and Kevin Gimpel	62
<i>Reading the Manual: Event Extraction as Definition Comprehension</i> Yunmo Chen, Tongfei Chen, Seth Ebner, Aaron Steven White and Benjamin Van Durme	74
<i>On the Discrepancy between Density Estimation and Sequence Generation</i> Jason Lee, Dustin Tran, Orhan Firat and Kyunghyun Cho	84
<i>Log-Linear Reformulation of the Noisy Channel Model for Document-Level Neural Machine Translation</i> Sébastien Jean and Kyunghyun Cho	95
<i>Deeply Embedded Knowledge Representation & Reasoning For Natural Language Question Answering: A Practitioner’s Perspective</i> Arindam Mitra, Sanjay Narayana and Chitta Baral	102

