EMNLP-IJCNLP 2019

Neural Generation and Translation

Proceedings of the Third Workshop

November 4, 2019
Hong Kong, China

©2019 The Association for Computational Linguistics

Apple and the Apple logo are trademarks of Apple Inc., registered in the U.S. and other countries.

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 78-1-950737-83-3

Introduction

Welcome to the Third Workshop on Neural Generation and Translation. This workshop aims to cultivate research on the leading edge in neural machine translation and other aspects of machine translation, generation, and multilinguality that utilize neural models. In this year's workshop we are extremely pleased to be able to host four invited talks from leading lights in the field, namely: Michael Auli, Mohit Bansal, Mirella Lapata and Jason Weston. In addition this year's workshop will feature a session devoted to a new shared task on efficient machine translation. We received a total of 68 submissions, from which we accepted 36. There were three crosssubmissions, seven long abstracts and 26 full papers. There were also seven system submission papers. All research papers were reviewed twice through a double blind review process, and avoiding conflicts of interest. The workshop had an acceptance rate of 53%. Due to the large number of invited talks, and to encourage discussion, only the two papers selected for best paper awards will be presented orally, and the remainder will be presented in a single poster session. We would like to thank all authors for their submissions, and the program committee members for their valuable efforts in reviewing the papers for the workshop. We would also like to thank Google and Apple for their generous sponsorship.

Organizers:

Alexandra Birch, (Edinburgh)
Andrew Finch, (Apple)
Hiroaki Hayashi (CMU)
Ioannis Konstas (Heriot Watt University)
Thang Luong, (Google)
Graham Neubig, (CMU)
Yusuke Oda, (Google)
Katsuhito Sudoh (NAIST)

Program Committee:

Roee Aharoni
Antonio Valerio Miceli Barone
Joost Bastings
Marine Carpuat
Daniel Cer
Boxing Chen
Eunah Cho
Li Dong
Nan Du
Kevin Duh
Ondřej Dušek
Markus Freitag
Claire Gardent
Ulrich Germann
Isao Goto
Edward Grefenstette
Roman Grundkiewicz
Jiatao Gu
Barry Haddow
Vu Cong Duy Hoang
Daphne Ippolito
Sébastien Jean
Hidetaka Kamigaito
Shumpei Kubosawa
Sneha Kudugunta
Lemao Liu
Shujie Liu
Hongyin Luo
Benjamin Marie
Sebastian J. Mielke
Hideya Mino
Makoto Morishita
Preslav Nakov
Vivek Natarajan
Laura Perez-Beltrachini
Vinay Rao
Alexander Rush
Chinnadhurai Sankar
Rico Sennrich
Raphael Shu
Akihiro Tamura
Rui Wang
Xiaolin Wang
Taro Watanabe
Sam Wiseman
Biao Zhang

Invited Speakers:

Michael Auli (Facebook AI Research)
Mohit Bansal (University of North Carolina)
Nanyun Peng (University of Southern California)
Jason Weston (Facebook AI Research)

Table of Contents

Findings of the Third Workshop on Neural Generation and Translation
Hiroaki Hayashi, Yusuke Oda, Alexandra Birch, Ioannis Konstas, Andrew Finch, Minh-Thang Luong, Graham Neubig and Katsuhito Sudoh 1

Hello, It's GPT-2 - How Can I Help You? Towards the Use of Pretrained Language Models for Task-Oriented Dialogue Systems
Paweł Budzianowski and Ivan Vulić 15

Recycling a Pre-trained BERT Encoder for Neural Machine Translation
Kenji Imamura and Eiichiro Sumita 23

Generating a Common Question from Multiple Documents using Multi-source Encoder-Decoder Models
Woon Sang Cho, Yizhe Zhang, Sudha Rao, Chris Brockett and Sungjin Lee 32

Generating Diverse Story Continuations with Controllable Semantics
Lifu Tu, Xiaoan Ding, Dong Yu and Kevin Gimpel 44

Domain Differential Adaptation for Neural Machine Translation
Zi-Yi Dou, Xinyi Wang, Junjie Hu and Graham Neubig 59

Transformer-based Model for Single Documents Neural Summarization
Elozino Egonmwan and Yllias Chali 70

Making Asynchronous Stochastic Gradient Descent Work for Transformers
Alham Fikri Aji and Kenneth Heafield 80

Controlled Text Generation for Data Augmentation in Intelligent Artificial Agents
Nikolaos Malandrakis, Minmin Shen, Anuj Goyal, Shuyang Gao, Abhishek Sethi and Angeliki Metallinou 90

Zero-Resource Neural Machine Translation with Monolingual Pivot Data
Anna Currey and Kenneth Heafield 99

On the use of BERT for Neural Machine Translation
Stephane Clinchant, Kweon Woo Jung and Vassilina Nikoulina 108

On the Importance of the Kullback-Leibler Divergence Term in Variational Autoencoders for Text Generation
Victor Prokhorov, Ehsan Shareghi, Yingzhen Li, Mohammad Taher Pilehvar and Nigel Collier 118

Decomposing Textual Information For Style Transfer
Ivan P. Yamshchikov, Viacheslav Shibaev, Aleksander Nagaev, Jürgen Jost and Alexey Tikhonov 128

Unsupervised Evaluation Metrics and Learning Criteria for Non-Parallel Textual Transfer
Richard Yuanzhe Pang and Kevin Gimpel 138

Enhanced Transformer Model for Data-to-Text Generation
Li GONG, Josep Crego and Jean Senellart 148

Generalization in Generation: A closer look at Exposure Bias
Florian Schmidt 157

Machine Translation of Restaurant Reviews: New Corpus for Domain Adaptation and Robustness
Alexandre Berard, Ioan Calapodescu, Marc Dymetman, Claude Roux, Jean-Luc Meunier and Vassilina Nikoulina 168

Adaptively Scheduled Multitask Learning: The Case of Low-Resource Neural Machine Translation
Poorya Zaremoodi and Gholamreza Haffari 177

On the Importance of Word Boundaries in Character-level Neural Machine Translation
Duygu Ataman, Orhan Firat, Mattia A. Di Gangi, Marcello Federico and Alexandra Birch 187

Big Bidirectional Insertion Representations for Documents
Lala Li and William Chan 194

A Margin-based Loss with Synthetic Negative Samples for Continuous-output Machine Translation
Gayatri Bhat, Sachin Kumar and Yulia Tsvetkov 199

Mixed Multi-Head Self-Attention for Neural Machine Translation
Hongyi Cui, Shohei Iida, Po-Hsuan Hung, Takehito Utsuro and Masaaki Nagata 206

Paraphrasing with Large Language Models
Sam Witteveen and Martin Andrews 215

Interrogating the Explanatory Power of Attention in Neural Machine Translation
Pooya Moradi, Nishant Kambhatla and Anoop Sarkar 221

Auto-Sizing the Transformer Network: Improving Speed, Efficiency, and Performance for Low-Resource Machine Translation
Kenton Murray, Jeffery Kinnison, Toan Q. Nguyen, Walter Scheirer and David Chiang 231

Learning to Generate Word- and Phrase-Embeddings for Efficient Phrase-Based Neural Machine Translation
Chan Young Park and Yulia Tsvetkov 241

Transformer and seq2seq model for Paraphrase Generation
Elozino Egonmwan and Yllias Chali 249

Monash University's Submissions to the WNGT 2019 Document Translation Task
Sameen Maruf and Gholamreza Haffari 256

SYSTRAN @ WNGT 2019: DGT Task
Li GONG, Josep Crego and Jean Senellart 262

University of Edinburgh's submission to the Document-level Generation and Translation Shared Task
Ratish Puduppully, Jonathan Mallinson and Mirella Lapata 268

Naver Labs Europe's Systems for the Document-Level Generation and Translation Task at WNGT 2019
Fahimeh Saleh, Alexandre Berard, Ioan Calapodescu and Laurent Besacier 273

From Research to Production and Back: Ludicrously Fast Neural Machine Translation
Young Jin Kim, Marcin Junczys-Dowmunt, Hany Hassan, Alham Fikri Aji, Kenneth Heafield, Roman Grundkiewicz and Nikolay Bogoychev 280

Selecting, Planning, and Rewriting: A Modular Approach for Data-to-Document Generation and Translation
Lesly Miculicich, Marc Marone and Hany Hassan 289

Efficiency through Auto-Sizing:Notre Dame NLP's Submission to the WNGT 2019 Efficiency Task
Kenton Murray, Brian DuSell and David Chiang 297

Workshop Program

09:00-09:10 *Welcome and Opening Remarks*
Findings of the Third Workshop on Neural Generation and Translation

09:10-10:00 *Keynote 1*
Michael Auli, Facebook AI Research

10:00-10:30 *Shared Task Overview*

10:30-10:40 *Shared Task Oral Presentation*

10:40-11:40 *Poster Session (see list below)*

11:40-12:30 *Keynote 2*
Jason Weston, Facebook AI Research

12:30-13:30 **Lunch Break**

13:30-14:20 *Keynote 3*
Nanyun Peng, USC

14:20-15:10 *Best Paper Session*

15:10-15:40 **Coffee Break**

15:20-16:30 *Keynote 4*
Mohit Bansal, University of North Carolina

16:30-17:00 **Closing Remarks**

Poster Session

Hello, It's GPT-2 - How Can I Help You? Towards the Use of Pretrained Language Models for Task-Oriented Dialogue Systems
Paweł Budzianowski and Ivan Vulić

Automated Generation of Search Advertisements
Jia Ying Jen, Divish Dayal, Corinne Choo, Ashish Awasthi, Audrey Kuah and Ziheng Lin

Recycling a Pre-trained BERT Encoder for Neural Machine Translation
Kenji Imamura and Eiichiro Sumita

Generating a Common Question from Multiple Documents using Multi-source Encoder-Decoder Models
Woon Sang Cho, Yizhe Zhang, Sudha Rao, Chris Brockett and Sungjin Lee

Positional Encoding to Control Output Sequence Length
Sho Takase and Naoaki Okazaki

Attending to Future Tokens for Bidirectional Sequence Generation
Carolin Lawrence, Bhushan Kotnis and Mathias Niepert

Generating Diverse Story Continuations with Controllable Semantics
Lifu Tu, Xiaoan Ding, Dong Yu and Kevin Gimpel

Domain Differential Adaptation for Neural Machine Translation
Zi-Yi Dou, Xinyi Wang, Junjie Hu and Graham Neubig

Transformer-based Model for Single Documents Neural Summarization
Elozino Egonmwan and Yllias Chali

Making Asynchronous Stochastic Gradient Descent Work for Transformers
Alham Fikri Aji and Kenneth Heafield

Controlled Text Generation for Data Augmentation in Intelligent Artificial Agents
Nikolaos Malandrakis, Minmin Shen, Anuj Goyal, Shuyang Gao, Abhishek Sethi and Angeliki Metallinou

Zero-Resource Neural Machine Translation with Monolingual Pivot Data
Anna Currey and Kenneth Heafield

November 4, 2019 (continued)

Two Birds, One Stone: A Simple, Unified Model for Text Generation from Structured and Unstructured Data
Hamidreza Shahidi, Ming Li and Jimmy Lin

On the use of BERT for Neural Machine Translation
Stephane Clinchant, Kweon Woo Jung and Vassilina Nikoulina

On the Importance of the Kullback-Leibler Divergence Term in Variational Autoencoders for Text Generation
Victor Prokhorov, Ehsan Shareghi, Yingzhen Li, Mohammad Taher Pilehvar and Nigel Collier

Decomposing Textual Information For Style Transfer
Ivan P. Yamshchikov, Viacheslav Shibaev, Aleksander Nagaev, Jürgen Jost and Alexey Tikhonov

Unsupervised Evaluation Metrics and Learning Criteria for Non-Parallel Textual Transfer
Richard Yuanzhe Pang and Kevin Gimpel

Enhanced Transformer Model for Data-to-Text Generation
Li GONG, Josep Crego and Jean Senellart

Generalization in Generation: A closer look at Exposure Bias
Florian Schmidt

Machine Translation of Restaurant Reviews: New Corpus for Domain Adaptation and Robustness
Alexandre Berard, Ioan Calapodescu, Marc Dymetman, Claude Roux, Jean-Luc Meunier and Vassilina Nikoulina

Improved Variational Neural Machine Translation via Promoting Mutual Information
Xian Li, Jiatao Gu, Ning Dong and Arya D. McCarthy

Adaptively Scheduled Multitask Learning: The Case of Low-Resource Neural Machine Translation
Poorya Zaremoodi and Gholamreza Haffari

Latent Relation Language Models
Hiroaki Hayashi, Zecong Hu, Chenyan Xiong and Graham Neubig

On the Importance of Word Boundaries in Character-level Neural Machine Translation
Duygu Ataman, Orhan Firat, Mattia A. Di Gangi, Marcello Federico and Alexandra Birch

November 4, 2019 (continued)

Big Bidirectional Insertion Representations for Documents
Lala Li and William Chan

The Daunting Task of Actual (Not Operational) Textual Style Transfer Auto-Evaluation
Richard Yuanzhe Pang

A Margin-based Loss with Synthetic Negative Samples for Continuous-output Machine Translation
Gayatri Bhat, Sachin Kumar and Yulia Tsvetkov

Context-Aware Learning for Neural Machine Translation
Sébastien Jean and Kyunghyun Cho

Mixed Multi-Head Self-Attention for Neural Machine Translation
Hongyi Cui, Shohei Iida, Po-Hsuan Hung, Takehito Utsuro and Masaaki Nagata

Paraphrasing with Large Language Models
Sam Witteveen and Martin Andrews

Interrogating the Explanatory Power of Attention in Neural Machine Translation
Pooya Moradi, Nishant Kambhatla and Anoop Sarkar

Insertion-Deletion Transformer
Laura Ruis, Mitchell Stern, Julia Proskurnia and William Chan

Auto-Sizing the Transformer Network: Improving Speed, Efficiency, and Performance for Low-Resource Machine Translation
Kenton Murray, Jeffery Kinnison, Toan Q. Nguyen, Walter Scheirer and David Chiang

Learning to Generate Word- and Phrase-Embeddings for Efficient Phrase-Based Neural Machine Translation
Chan Young Park and Yulia Tsvetkov

Transformer and seq2seq model for Paraphrase Generation
Elozino Egonmwan and Yllias Chali

Multilingual KERMIT: It's Not Easy Being Generative
Harris Chan, Jamie Kiros and William Chan

November 4, 2019 (continued)

Monash University's Submissions to the WNGT 2019 Document Translation Task
Sameen Maruf and Gholamreza Haffari

SYSTRAN @ WNGT 2019: DGT Task
Li GONG, Josep Crego and Jean Senellart

University of Edinburgh's submission to the Document-level Generation and Translation Shared Task
Ratish Puduppully, Jonathan Mallinson and Mirella Lapata

Naver Labs Europe's Systems for the Document-Level Generation and Translation Task at WNGT 2019
Fahimeh Saleh, Alexandre Berard, Ioan Calapodescu and Laurent Besacier

From Research to Production and Back: Ludicrously Fast Neural Machine Translation
Young Jin Kim, Marcin Junczys-Dowmunt, Hany Hassan, Alham Fikri Aji, Kenneth Heafield, Roman Grundkiewicz and Nikolay Bogoychev

Selecting, Planning, and Rewriting: A Modular Approach for Data-to-Document Generation and Translation
Lesly Miculicich, Marc Marone and Hany Hassan

Auto-Sizing the Transformer Network: Shrinking Parameters for the WNGT 2019 Efficiency Task
Kenton Murray, Brian DuSell and David Chiang