The 13th Conference of
The Association for Machine Translation
in the Americas

www.conference.amtaweb.org

PROCEEDINGS

Vol. 1: MT Researchers' Track

Editors: Colin Cherry & Graham Neubig

©2018 The Authors. These articles are licensed under a Creative Commons 3.0 license, no derivative works, attribution, CC-BY-ND.
Introduction

This year’s research track proceedings reflect AMTA’s long-standing tradition of bringing together machine translation researchers, practitioners and users. We invited ten long papers and four short papers for oral presentation, with the authors coming from around the world, and the papers spanning a diverse range of topics. These include improvements to the now-dominant neural machine translation paradigm, which address its lexical accuracy, its handling of morphology and its use of wider context. Meanwhile, we see papers that re-examine machine translation under specific lenses, studying its ability to maintain sentiment, to resist noisy inputs, and to retain user trust.

We have also introduced a new category to the research track proceedings: the Open Source Showcase, with the intent of bringing together the machine translation open source software community, and to highlight their amazing work to the rest of the AMTA community. This year, we are excited to present six open source system descriptions.

It has been a pleasure working with Olga Beregovaya and the rest of the AMTA 2018 organizing committee. We would like to especially thank steering committee members George Foster and Alon Lavie for their advice and support, as well as all the authors and reviewers, whose hard work made this conference possible. Please enjoy this year’s AMTA research track.

Sincerely,

Colin Cherry
Graham Neubig
Research Program Committee

Tim Anderson  Air Force Research Laboratory
Amittai Axelrod  Amazon
Marine Carpuat  University of Maryland
Daniel Cer  Google
Boxing Chen  Alibaba
David Chiang  University of Notre Dame
Nadir Durrani  QCRI
Minwei Feng  IBM Watson Group
George Foster  Google
Yvette Graham  Dublin City University
Yifan He  Bosch Research and Technology Center
Ulf Hermjakob  USC Information Sciences Institute
Matthias Huck  LMU Munich
Marcin Junczys-Dowmunt  Microsoft Research
Philipp Koehn  Johns Hopkins University
Roland Kuhn  National Research Council of Canada
Gregor Leusch  eBay
Lemao Liu  Tencent AI Lab
Qun Liu  Dublin City University
Saab Mansour  Apple
Daniel Marcu  USC Information Sciences Institute
Haitao Mi  Alipay US
Rico Sennrich  University of Edinburgh
Michel Simard  National Research Council of Canada
Raymond Slyh  Air Force Research Laboratory
Jörg Tiedemann  University of Helsinki
Christoph Tillmann  IBM Research
Taro Watanabe  Google
Philip Williams  University of Edinburgh
François Yvon  LIMSI/CNRS
Bing Zhao  SRI International
## Contents

1 Document-Level Information as Side Constraints for Improved Neural Patent Translation  
Laura Jehl and Stefan Riezler  

13 Fluency Over Adequacy: A Pilot Study in Measuring User Trust in Imperfect MT  
Marianna J. Martindale and Marine Carpuat  

26 Combining Quality Estimation and Automatic Post-editing to Enhance Machine Translation output  
Rajen Chatterjee, Matteo Negri, Marco Turchi, Frédéric Blain and Lucia Specia  

39 Neural Morphological Tagging of Lemma Sequences for Machine Translation  
Costanza Conforti, Matthias Huck and Alexander Fraser  

54 Context Models for OOV Word Translation in Low-Resource Languages  
Angli Liu and Katrin Kirchhoff  

68 How Robust Are Character-Based Word Embeddings in Tagging and MT Against Word Scrambling or Random Noise?  
Georg Heigold, Stalin Varanasi, Günter Neumann and Josef van Genabith  

81 Balancing Translation Quality and Sentiment Preservation (Non-archival Extended Abstract)  
Pintu Lohar, Haithem Afli and Andy Way
89  Register-sensitive Translation: a Case Study of Mandarin and Cantonese (Non-archival Extended Abstract)

Tak-sum Wong and John Lee

97  An Evaluation of Two Vocabulary Reduction Methods for Neural Machine Translation

Duygu Ataman and Marcello Federico

111 A Smorgasbord of Features to Combine Phrase-Based and Neural Machine Translation

Benjamin Marie and Atsushi Fujita

125 Exploring Word Sense Disambiguation Abilities of Neural Machine Translation Systems (Non-archival Extended Abstract)

Rebecca Marvin and Philipp Koehn

132 Improving Low Resource Machine Translation using Morphological Glosses (Non-archival Extended Abstract)

Steven Shearing, Christo Kirov, Huda Khayrallah and David Yarowsky

140 A Dataset and Reranking Method for Multimodal MT of User-Generated Image Captions

Shigehiko Schamoni, Julian Hitschler and Stefan Riezler

154 Simultaneous Translation using Optimized Segmentation

Maryam Siahbani, Hassan S. Shavarani, Ashkan Alinejad and Anoop Sarkar
Open Source Showcase

168  Neural Monkey: The Current State and Beyond
    Jindřich Helcl, Jindřich Libovický, Tom Kocmi, Tomáš Musil, Ondřej Cífka, Dusan Varis and Ondřej Bojar

177  OpenNMT: Neural Machine Translation Toolkit
    Guillaume Klein, Yoon Kim, Yuntian Deng, Vincent Nguyen, Jean Senellart and Alexander Rush

185  XNMT: The eXtensible Neural Machine Translation Toolkit

193  Tensor2Tensor for Neural Machine Translation

200  The Sockeye Neural Machine Translation Toolkit at AMTA 2018
    Felix Hieber, Tobias Domhan, Michael Denkowski, David Vilar, Artem Sokolov, Ann Clifton and Matt Post

208  Why not be Versatile? Applications of the SGNMT Decoder for Machine Translation
    Felix Stahlberg, Danielle Saunders, Gonzalo Iglesias and Bill Byrne