

9th Workshop on Statistical Machine Translation 2014

Held at ACL 2014

**Baltimore, Maryland, USA
26 – 27 June 2014**

ISBN: 978-1-63439-213-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2014) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Efficient Elicitation of Annotations for Human Evaluation of Machine Translation</i> Keisuke Sakaguchi, Matt Post and Benjamin Van Durme	1
<i>Findings of the 2014 Workshop on Statistical Machine Translation</i> Ondrej Bojar, Christian Buck, Christian Federmann, Barry Haddow, Philipp Koehn, Johannes Leve- ling, Christof Monz, Pavel Pecina, Matt Post, Herve Saint-Amand, Radu Soricut, Lucia Specia and Aleš Tamchyna	12
<i>Parallel FDA5 for Fast Deployment of Accurate Statistical Machine Translation Systems</i> Ergun Bicici, Qun Liu and Andy Way	59
<i>Yandex School of Data Analysis Russian-English Machine Translation System for WMT14</i> Alexey Borisov and Irina Galinskaya	66
<i>CimS – The CIS and IMS joint submission to WMT 2014 translating from English into German</i> Fabienne Cap, Marion Weller, Anita Ramm and Alexander Fraser	71
<i>English-to-Hindi system description for WMT 2014: Deep Source-Context Features for Moses</i> Marta R. Costa-jussà, Parth Gupta, Paolo Rosso and Rafael E. Banchs	79
<i>The KIT-LIMSI Translation System for WMT 2014</i> Quoc Khanh Do, Teresa Herrmann, Jan Niehues, Alexander Allauzen, François Yvon and Alex Waibel	84
<i>The IIT Bombay Hindi-English Translation System at WMT 2014</i> Piyush Dungarwal, Rajen Chatterjee, Abhijit Mishra, Anoop Kunchukuttan, Ritesh Shah and Push- pak Bhattacharyya	90
<i>Edinburgh’s Phrase-based Machine Translation Systems for WMT-14</i> Nadir Durrani, Barry Haddow, Philipp Koehn and Kenneth Heafield	97
<i>EU-BRIDGE MT: Combined Machine Translation</i> Markus Freitag, Stephan Peitz, Joern Wuebker, Hermann Ney, Matthias Huck, Rico Sennrich, Nadir Durrani, Maria Nadejde, Philip Williams, Philipp Koehn, Teresa Herrmann, Eunah Cho and Alex Waibel 105	
<i>Phrasal: A Toolkit for New Directions in Statistical Machine Translation</i> Spence Green, Daniel Cer and Christopher Manning	114
<i>Anaphora Models and Reordering for Phrase-Based SMT</i> Christian Hardmeier, Sara Stymne, Jörg Tiedemann, Aaron Smith and Joakim Nivre	122
<i>The Karlsruhe Institute of Technology Translation Systems for the WMT 2014</i> Teresa Herrmann, Mohammed Mediani, Eunah Cho, Thanh-Le Ha, Jan Niehues, Isabel Slawik, Yuqi Zhang and Alex Waibel	130
<i>The DCU-ICTCAS MT system at WMT 2014 on German-English Translation Task</i> Liangyou Li, Xiaofeng Wu, Santiago Cortes Vaillo, Jun Xie, Andy Way and Qun Liu	136
<i>The CMU Machine Translation Systems at WMT 2014</i> Austin Matthews, Waleed Ammar, Archana Bhatia, Weston Feely, Greg Hanneman, Eva Schlinger, Swabha Swayamdipta, Yulia Tsvetkov, Alon Lavie and Chris Dyer	142

<i>Stanford University's Submissions to the WMT 2014 Translation Task</i>	
Julia Neidert, Sebastian Schuster, Spence Green, Kenneth Heafield and Christopher Manning .	150
<i>The RWTH Aachen German-English Machine Translation System for WMT 2014</i>	
Stephan Peitz, Joern Wuebker, Markus Freitag and Hermann Ney	157
<i>Large-scale Exact Decoding: The IMS-TTT submission to WMT14</i>	
Daniel Quernheim and Fabienne Cap	163
<i>Abu-MaTran at WMT 2014 Translation Task: Two-step Data Selection and RBMT-Style Synthetic Rules</i>	
Raphael Rubino, Antonio Toral, Víctor M. Sánchez-Cartagena, Jorge Ferrández-Tordera, Sergio Ortiz Rojas, Gema Ramírez-Sánchez, Felipe Sánchez-Martínez and Andy Way	171
<i>The UA-Prompsit hybrid machine translation system for the 2014 Workshop on Statistical Machine Translation</i>	
Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez	178
<i>Machine Translation and Monolingual Postediting: The AFRL WMT-14 System</i>	
Lane Schwartz, Timothy Anderson, Jeremy Gwinnup and Katherine Young	186
<i>CUNI in WMT14: Chimera Still Awaits Bellerophon</i>	
Aleš Tamchyna, Martin Popel, Rudolf Rosa and Ondrej Bojar	195
<i>Manawi: Using Multi-Word Expressions and Named Entities to Improve Machine Translation</i>	
Liling Tan and Santanu Pal	201
<i>Edinburgh's Syntax-Based Systems at WMT 2014</i>	
Philip Williams, Rico Sennrich, Maria Nadejde, Matthias Huck, Eva Hasler and Philipp Koehn	207
<i>DCU-Lingo24 Participation in WMT 2014 Hindi-English Translation task</i>	
Xiaofeng Wu, Rejwanul Haque, Tsuyoshi Okita, Piyush Arora, Andy Way and Qun Liu	215
<i>Machine Translation of Medical Texts in the Khresmoi Project</i>	
Ondřej Dušek, Jan Hajič, Jaroslava Hlaváčová, Michal Novák, Pavel Pecina, Rudolf Rosa, Aleš Tamchyna, Zdeňka Urešová and Daniel Zeman	221
<i>Postech's System Description for Medical Text Translation Task</i>	
Jianri Li, Se-Jong Kim, Hwidong Na and Jong-Hyeok Lee	229
<i>Domain Adaptation for Medical Text Translation using Web Resources</i>	
Yi Lu, Longyue Wang, Derek F. Wong, Lidia S. Chao and Yiming Wang	233
<i>DCU Terminology Translation System for Medical Query Subtask at WMT14</i>	
Tsuyoshi Okita, Ali Vahid, Andy Way and Qun Liu	239
<i>LIMSI @ WMT'14 Medical Translation Task</i>	
Nicolas Pécheux, Li Gong, Quoc Khanh Do, Benjamin Marie, Yulia Ivanishcheva, Alexander Al-lauzen, Thomas Lavergne, Jan Niehues, Aurélien Max and François Yvon	246
<i>Combining Domain Adaptation Approaches for Medical Text Translation</i>	
Longyue Wang, Yi Lu, Derek F. Wong, Lidia S. Chao, Yiming Wang and Francisco Oliveira . .	254
<i>Experiments in Medical Translation Shared Task at WMT 2014</i>	
Jian Zhang	260

<i>Randomized Significance Tests in Machine Translation</i>	
Yvette Graham, Nitika Mathur and Timothy Baldwin	266
<i>Estimating Word Alignment Quality for SMT Reordering Tasks</i>	
Sara Stymne, Jörg Tiedemann and Joakim Nivre	275
<i>Dependency-based Automatic Enumeration of Semantically Equivalent Word Orders for Evaluating Japanese Translations</i>	
Hideki Isozaki, Natsume Kouchi and Tsutomu Hirao	287
<i>Results of the WMT14 Metrics Shared Task</i>	
Matous Machacek and Ondrej Bojar	293
<i>Efforts on Machine Learning over Human-mediated Translation Edit Rate</i>	
Eleftherios Avramidis	302
<i>SHEF-Lite 2.0: Sparse Multi-task Gaussian Processes for Translation Quality Estimation</i>	
Daniel Beck, Kashif Shah and Lucia Specia	307
<i>Referential Translation Machines for Predicting Translation Quality</i>	
Ergun Bici and Andy Way	313
<i>FBK-UPV-UEdin participation in the WMT14 Quality Estimation shared-task</i>	
José Guilherme Camargo de Souza, Jesús González-Rubio, Christian Buck, Marco Turchi and Matteo Negri	322
<i>Target-Centric Features for Translation Quality Estimation</i>	
Chris Hokamp, Iacer Calixto, Joachim Wagner and Jian Zhang	329
<i>LIG System for Word Level QE task at WMT14</i>	
Ngoc Quang Luong, Laurent Besacier and Benjamin Lecouteux	335
<i>Exploring Consensus in Machine Translation for Quality Estimation</i>	
Carolina Scarton and Lucia Specia	342
<i>LIMSI Submission for WMT'14 QE Task</i>	
Guillaume Wisniewski, Nicolas Pécheux, Alexander Allauzen and François Yvon	348
<i>Parmesan: Meteor without Paraphrases with Paraphrased References</i>	
Petra Barancikova	355
<i>A Systematic Comparison of Smoothing Techniques for Sentence-Level BLEU</i>	
Boxing Chen and Colin Cherry	362
<i>VERTa participation in the WMT14 Metrics Task</i>	
Elisabet Comelles and Jordi Atserias	368
<i>Meteor Universal: Language Specific Translation Evaluation for Any Target Language</i>	
Michael Denkowski and Alon Lavie	376
<i>Application of Prize based on Sentence Length in Chunk-based Automatic Evaluation of Machine Translation</i>	
Hiroshi Echizen'ya, Kenji Araki and Eduard Hovy	381
<i>LAYERED: Metric for Machine Translation Evaluation</i>	
Shubham Gautam and Pushpak Bhattacharyya	387

<i>IPA and STOUT: Leveraging Linguistic and Source-based Features for Machine Translation Evaluation</i> Meritxell González, Alberto Barrón-Cedeño and Lluís Màrquez	394
<i>DiscoTK: Using Discourse Structure for Machine Translation Evaluation</i> Shafiq Joty, Francisco Guzmán, Lluís Màrquez and Preslav Nakov	402
<i>Tolerant BLEU: a Submission to the WMT14 Metrics Task</i> Jindřich Libovický and Pavel Pecina	409
<i>BEER: BEtter Evaluation as Ranking</i> Milos Stanojevic and Khalil Sima'an	414
<i>RED, The DCU-CASICT Submission of Metrics Tasks</i> Xiaofeng Wu, Hui Yu and Qun Liu	420
<i>Crowdsourcing High-Quality Parallel Data Extraction from Twitter</i> Wang Ling, Luis Marujo, Chris Dyer, Alan W Black and Isabel Trancoso	426
<i>Using Comparable Corpora to Adapt MT Models to New Domains</i> Ann Irvine and Chris Callison-Burch	437
<i>Dynamic Topic Adaptation for SMT using Distributional Profiles</i> Eva Hasler, Barry Haddow and Philipp Koehn	445
<i>Unsupervised Adaptation for Statistical Machine Translation</i> Saab Mansour and Hermann Ney	457
<i>An Empirical Comparison of Features and Tuning for Phrase-based Machine Translation</i> Spence Green, Daniel Cer and Christopher Manning	466
<i>Bayesian Reordering Model with Feature Selection</i> Abdullah Alrajeh and Mahesan Niranjan	477
<i>Augmenting String-to-Tree and Tree-to-String Translation with Non-Syntactic Phrases</i> Matthias Huck, Hieu Hoang and Philipp Koehn	486
<i>Linear Mixture Models for Robust Machine Translation</i> Marine Carpuat, Cyril Goutte and George Foster	499