ACL 2014

Ninth Workshop on Statistical Machine Translation

Proceedings of the Workshop

June 26-27, 2014
Baltimore, Maryland, USA

©2014 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 978-1-941643-17-4

Introduction

The ACL 2014 Workshop on Statistical Machine Translation (WMT 2014) took place on Thursday and Friday, June 26–27, 2014 in Baltimore, United States, immediately following the Conference of the Association for Computational Linguistics (ACL).

This is the ninth time this workshop has been held. The first time it was held at HLT-NAACL 2006 in New York City, USA. In the following years the Workshop on Statistical Machine Translation was held at ACL 2007 in Prague, Czech Republic, ACL 2008, Columbus, Ohio, USA, EACL 2009 in Athens, Greece, ACL 2010 in Uppsala, Sweden, EMNLP 2011 in Edinburgh, Scotland, NAACL 2012 in Montreal, Canada, and ACL 2013 in Sofia, Bulgaria.

The focus of our workshop was to use parallel corpora for machine translation. Recent experimentation has shown that the performance of SMT systems varies greatly with the source language. In this workshop we encouraged researchers to investigate ways to improve the performance of SMT systems for diverse languages, including morphologically more complex languages, languages with partial free word order, and low-resource languages.

Prior to the workshop, in addition to soliciting relevant papers for review and possible presentation, we conducted four shared tasks: a general translation task, a medical translation task, a quality estimation task, and a task to test automatic evaluation metrics. The medical translation task was introduced this year to address the important issue of domain adaptation within SMT. The results of the shared tasks were announced at the workshop, and these proceedings also include an overview paper for the shared tasks that summarizes the results, as well as provides information about the data used and any procedures that were followed in conducting or scoring the task. In addition, there are short papers from each participating team that describe their underlying system in greater detail.

Like in previous years, we have received a far larger number of submission than we could accept for presentation. This year we have received 27 full paper submissions and 49 shared task submissions. In total WMT 2014 featured 12 full paper oral presentations and 49 shared task poster presentations.

The invited talk was given by Alon Lavie (Carnegie Mellon University and Safaba Translation Solutions, Inc.) entitled "Machine Translation in Academia and in the Commercial World – a Contrastive Perspective".

We would like to thank the members of the Program Committee for their timely reviews. We also would like to thank the participants of the shared task and all the other volunteers who helped with the evaluations.

Ondřej Bojar, Christian Buck, Christian Federmann, Barry Haddow, Philipp Koehn, Matouš Macháček, Christof Monz, Pavel Pecina, Matt Post, Hervé Saint-Amand, Radu Soricut, and Lucia Specia

Co-Organizers

Organizers:

Ondřej Bojar (Charles University Prague)
Christian Buck (University of Edinburgh)
Christian Federmann (Microsoft Research)
Barry Haddow (University of Edinburgh)
Philipp Koehn (University of Edinburgh / Johns Hopkins University)
Matouš Macháček (Charles University Prague)
Christof Monz (University of Amsterdam)
Pavel Pecina (Charles University Prague)
Matt Post (Johns Hopkins University)
Hervé Saint-Amand (University of Edinburgh)
Radu Soricut (Google)
Lucia Specia (University of Sheffield)

Invited Talk:

Alon Lavie (Research Professor at Carnegie Mellon University / Co-founder, President and CTO - Safaba Translation Solutions, Inc.)

Program Committee:

Lars Ahrenberg (Linköping University)
Alexander Allauzen (Université Paris-Sud / LIMSI-CNRS)
Tim Anderson (Air Force Research Laboratory)
Eleftherios Avramidis (German Research Center for Artificial Intelligence)
Wilker Aziz (University of Sheffield)
Daniel Beck (University of Sheffield)
Jose Miguel Benedi (Universitàt Politecnica de València)
Nicola Bertoldi (FBK)
Ergun Bicici (Centre for Next Generation Localisation, Dublin City University)
Alexandra Birch (University of Edinburgh)
Arianna Bisazza (University of Amsterdam)
Graeme Blackwood (IBM Research)
Phil Blunsom (University of Oxford)
Fabienne Braune (University of Stuttgart)
Chris Brockett (Microsoft Research)
Hailong Cao (Harbin Institute of Technology)
Michael Carl (Copenhagen Business School)
Marine Carpuat (National Research Council)
Francisco Casacuberta (Universitat Politécnica de Valéncia)
Daniel Cer (Google)
Boxing Chen (NRC)
Colin Cherry (NRC)
David Chiang (USC/ISI)
Vishal Chowdhary (Microsoft)

Steve DeNeefe (SDL Language Weaver)
Michael Denkowski (Carnegie Mellon University)
Jacob Devlin (Raytheon BBN Technologies)
Markus Dreyer (SDL Language Weaver)
Kevin Duh (Nara Institute of Science and Technology)
Marcello Federico (FBK)
Yang Feng (USC/ISI)
Andrew Finch (NICT)
Mark Fishel (University of Zurich)
José A. R. Fonollosa (Universitat Politècnica de Catalunya)
George Foster (NRC)
Michel Galley (Microsoft Research)
Juri Ganitkevitch (Johns Hopkins University)
Katya Garmash (University of Amsterdam)
Josef van Genabith (Dublin City University)
Ulrich Germann (University of Edinburgh)
Daniel Gildea (University of Rochester)
Kevin Gimpel (Toyota Technological Institute at Chicago)
Jesús González-Rubio (Universitat Politécnica de Valéncia)
Yvette Graham (The University of Melbourne)
Spence Green (Stanford University)
Francisco Guzmán (Qatar Computing Research Institute)
Greg Hanneman (Carnegie Mellon University)
Christian Hardmeier (Uppsala universitet)
Eva Hasler (University of Edinburgh)
Yifan He (New York University)
Kenneth Heafield (Stanford)
John Henderson (MITRE)
Felix Hieber (Heidelberg University)
Hieu Hoang (University of Edinburgh)
Stéphane Huet (UniversitÃ© d'Avignon)
Young-Sook Hwang (SKPlanet)
Gonzalo Iglesias (University of Cambridge)
Ann Irvine (Johns Hopkins University)
Abe Ittycheriah (IBM)
Laura Jehl (Heidelberg University)
Doug Jones (MIT Lincoln Laboratory)
Maxim Khalilov (BMMT)
Alexander Koller (University of Potsdam)
Roland Kuhn (National Research Council of Canada)
Shankar Kumar (Google)
Mathias Lambert (Amazon.com)
Phillippe Langlais (Université de Montréal)
Alon Lavie (Carnegie Mellon University)
Gennadi Lembersky (NICE Systems)
William Lewis (Microsoft Research)
Lemao Liu (The City University of New York)

Qun Liu (Dublin City University)
Wolfgang Macherey (Google)
Saab Mansour (RWTH Aachen University)
José B. Mariño (Universitat Politècnica de Catalunya)
Cettolo Mauro (FBK)
Arne Mauser (Google, Inc)
Jon May (SDL Language Weaver)
Wolfgang Menzel (Hamburg University)
Shachar Mirkin (Xerox Research Centre Europe)
Yusuke Miyao (National Instutite of Informatics)
Dragos Munteanu (SDL Language Technologies)
Markos Mylonakis (Lexis Research)
Lluís Màrquez (Qatar Computing Research Institute)
Preslav Nakov (Qatar Computing Research Institute)
Graham Neubig (Nara Institute of Science and Technology)
Jan Niehues (Karlsruhe Institute of Technology)
Kemal Oflazer (Carnegie Mellon University - Qatar)
Daniel Ortiz-Martínez (Copenhagen Business School)
Stephan Peitz (RWTH Aachen University)
Sergio Penkale (Lingo24)
Maja Popović (DFKI)
Stefan Riezler (Heidelberg University)
Johann Roturier (Symantec)
Raphael Rubino (Prompsit Language Engineering)
Alexander M. Rush (MIT)
Anoop Sarkar (Simon Fraser University)
Hassan Sawaf (eBay Inc.)
Lane Schwartz (Air Force Research Laboratory)
Jean Senellart (SYSTRAN)
Rico Sennrich (University of Zurich)
Kashif Shah (University of Sheffield)
Wade Shen (MIT)
Patrick Simianer (Heidelberg University)
Linfeng Song (ICT/CAS)
Sara Stymne (Uppsala University)
Katsuhito Sudoh (NTT Communication Science Laboratories / Kyoto University)
Felipe Sánchez-Martínez (Universitat d'Alacant)
Jörg Tiedemann (Uppsala University)
Christoph Tillmann (TJ Watson IBM Research)
Antonio Toral (Dublin City Unversity)
Hajime Tsukada (NTT Communication Science Laboratories)
Yulia Tsvetkov (Carnegie Mellon University)
Dan Tufiş (Research Institute for Artificial Intelligence, Romanian Academy)
Marco Turchi (Fondazione Bruno Kessler)
Ferhan Ture (University of Maryland)
Masao Utiyama (NICT)
Ashish Vaswani (University of Southern California Information Sciences Institute)

David Vilar (Pixformance GmbH)
Stephan Vogel (Qatar Computing Research Institute)
Haifeng Wang (Baidu)
Taro Watanabe (NICT)
Marion Weller (Universität Stuttgart)
Philip Williams (University of Edinburgh)
Guillaume Wisniewski (Univ. Paris Sud and LIMSI-CNRS)
Hua Wu (Baidu)
Joern Wuebker (RWTH Aachen University)
Peng Xu (Google Inc.)
Wenduan Xu (Cambridge University)
François Yvon (LIMSI/CNRS)
Richard Zens (Google)
Hao Zhang (Google)
Liu Zhanyi (Baidu)

Table of Contents

Efficient Elicitation of Annotations for Human Evaluation of Machine Translation
Keisuke Sakaguchi, Matt Post and Benjamin Van Durme 1

Findings of the 2014 Workshop on Statistical Machine Translation
Ondrej Bojar, Christian Buck, Christian Federmann, Barry Haddow, Philipp Koehn, Johannes Leveling, Christof Monz, Pavel Pecina, Matt Post, Herve Saint-Amand, Radu Soricut, Lucia Specia and Aleš Tamchyna 12

Parallel FDA5 for Fast Deployment of Accurate Statistical Machine Translation Systems
Ergun Bicici, Qun Liu and Andy Way 59

Yandex School of Data Analysis Russian-English Machine Translation System for WMT14
Alexey Borisov and Irina Galinskaya 66

CimS – The CIS and IMS joint submission to WMT 2014 translating from English into German
Fabienne Cap, Marion Weller, Anita Ramm and Alexander Fraser 71

English-to-Hindi system description for WMT 2014: Deep Source-Context Features for Moses
Marta R. Costa-jussà, Parth Gupta, Paolo Rosso and Rafael E. Banchs 79

The KIT-LIMSI Translation System for WMT 2014
Quoc Khanh Do, Teresa Herrmann, Jan Niehues, Alexander Allauzen, François Yvon and Alex Waibel 84

The IIT Bombay Hindi-English Translation System at WMT 2014
Piyush Dungarwal, Rajen Chatterjee, Abhijit Mishra, Anoop Kunchukuttan, Ritesh Shah and Pushpak Bhattacharyya 90

Edinburgh's Phrase-based Machine Translation Systems for WMT-14
Nadir Durrani, Barry Haddow, Philipp Koehn and Kenneth Heafield 97

EU-BRIDGE MT: Combined Machine Translation
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

Phrasal: A Toolkit for New Directions in Statistical Machine Translation
Spence Green, Daniel Cer and Christopher Manning 114

Anaphora Models and Reordering for Phrase-Based SMT
Christian Hardmeier, Sara Stymne, Jörg Tiedemann, Aaron Smith and Joakim Nivre 122

The Karlsruhe Institute of Technology Translation Systems for the WMT 2014
Teresa Herrmann, Mohammed Mediani, Eunah Cho, Thanh-Le Ha, Jan Niehues, Isabel Slawik, Yuqi Zhang and Alex Waibel 130

The DCU-ICTCAS MT system at WMT 2014 on German-English Translation Task
Liangyou Li, Xiaofeng Wu, Santiago Cortes Vaillo, Jun Xie, Andy Way and Qun Liu 136

The CMU Machine Translation Systems at WMT 2014
Austin Matthews, Waleed Ammar, Archna Bhatia, Weston Feely, Greg Hanneman, Eva Schlinger, Swabha Swayamdipta, Yulia Tsvetkov, Alon Lavie and Chris Dyer 142

Stanford University's Submissions to the WMT 2014 Translation Task
Julia Neidert, Sebastian Schuster, Spence Green, Kenneth Heafield and Christopher Manning . 150

The RWTH Aachen German-English Machine Translation System for WMT 2014
Stephan Peitz, Joern Wuebker, Markus Freitag and Hermann Ney 157

Large-scale Exact Decoding: The IMS-TTT submission to WMT14
Daniel Quernheim and Fabienne Cap 163

Abu-MaTran at WMT 2014 Translation Task: Two-step Data Selection and RBMT-Style Synthetic Rules
Raphael Rubino, Antonio Toral, Víctor M. Sánchez-Cartagena, Jorge Ferrández-Tordera, Sergio Ortiz Rojas, Gema Ramírez-Sánchez, Felipe Sánchez-Martínez and Andy Way 171

The UA-Prompsit hybrid machine translation system for the 2014 Workshop on Statistical Machine Translation
Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez 178

Machine Translation and Monolingual Postediting: The AFRL WMT-14 System
Lane Schwartz, Timothy Anderson, Jeremy Gwinnup and Katherine Young 186

CUNI in WMT14: Chimera Still Awaits Bellerophon
Aleš Tamchyna, Martin Popel, Rudolf Rosa and Ondrej Bojar 195

Manawi: Using Multi-Word Expressions and Named Entities to Improve Machine Translation
Liling Tan and Santanu Pal 201

Edinburgh's Syntax-Based Systems at WMT 2014
Philip Williams, Rico Sennrich, Maria Nadejde, Matthias Huck, Eva Hasler and Philipp Koehn 207

DCU-Lingo24 Participation in WMT 2014 Hindi-English Translation task
Xiaofeng Wu, Rejwanul Haque, Tsuyoshi Okita, Piyush Arora, Andy Way and Qun Liu 215

Machine Translation of Medical Texts in the Khresmoi Project
Ondřej Dušek, Jan Hajič, Jaroslava Hlaváčová, Michal Novák, Pavel Pecina, Rudolf Rosa, Aleš Tamchyna, Zdeňka Urešová and Daniel Zeman 221

Postech's System Description for Medical Text Translation Task
Jianri Li, Se-Jong Kim, Hwidong Na and Jong-Hyeok Lee 229

Domain Adaptation for Medical Text Translation using Web Resources
Yi Lu, Longyue Wang, Derek F. Wong, Lidia S. Chao and Yiming Wang 233

DCU Terminology Translation System for Medical Query Subtask at WMT14
Tsuyoshi Okita, Ali Vahid, Andy Way and Qun Liu 239

LIMSI @ WMT'14 Medical Translation Task
Nicolas Pécheux, Li Gong, Quoc Khanh Do, Benjamin Marie, Yulia Ivanishcheva, Alexander Allauzen, Thomas Lavergne, Jan Niehues, Aurélien Max and François Yvon 246

Combining Domain Adaptation Approaches for Medical Text Translation
Longyue Wang, Yi Lu, Derek F. Wong, Lidia S. Chao, Yiming Wang and Francisco Oliveira . . 254

Experiments in Medical Translation Shared Task at WMT 2014
Jian Zhang 260

Randomized Significance Tests in Machine Translation
Yvette Graham, Nitika Mathur and Timothy Baldwin 266

Estimating Word Alignment Quality for SMT Reordering Tasks
Sara Stymne, Jörg Tiedemann and Joakim Nivre 275

Dependency-based Automatic Enumeration of Semantically Equivalent Word Orders for Evaluating Japanese Translations
Hideki Isozaki, Natsume Kouchi and Tsutomu Hirao 287

Results of the WMT14 Metrics Shared Task
Matous Machacek and Ondrej Bojar 293

Efforts on Machine Learning over Human-mediated Translation Edit Rate
Eleftherios Avramidis 302

SHEF-Lite 2.0: Sparse Multi-task Gaussian Processes for Translation Quality Estimation
Daniel Beck, Kashif Shah and Lucia Specia 307

Referential Translation Machines for Predicting Translation Quality
Ergun Bicici and Andy Way 313

FBK-UPV-UEdin participation in the WMT14 Quality Estimation shared-task
José Guilherme Camargo de Souza, Jesús González-Rubio, Christian Buck, Marco Turchi and Matteo Negri 322

Target-Centric Features for Translation Quality Estimation
Chris Hokamp, Iacer Calixto, Joachim Wagner and Jian Zhang 329

LIG System for Word Level QE task at WMT14
Ngoc Quang Luong, Laurent Besacier and Benjamin Lecouteux 335

Exploring Consensus in Machine Translation for Quality Estimation
Carolina Scarton and Lucia Specia 342

LIMSI Submission for WMT'14 QE Task
Guillaume Wisniewski, Nicolas Pécheux, Alexander Allauzen and François Yvon 348

Parmesan: Meteor without Paraphrases with Paraphrased References
Petra Barancikova 355

A Systematic Comparison of Smoothing Techniques for Sentence-Level BLEU
Boxing Chen and Colin Cherry 362

VERTa participation in the WMT14 Metrics Task
Elisabet Comelles and Jordi Atserias 368

Meteor Universal: Language Specific Translation Evaluation for Any Target Language
Michael Denkowski and Alon Lavie 376

Application of Prize based on Sentence Length in Chunk-based Automatic Evaluation of Machine Translation
Hiroshi Echizen'ya, Kenji Araki and Eduard Hovy 381

LAYERED: Metric for Machine Translation Evaluation
Shubham Gautam and Pushpak Bhattacharyya 387

IPA and STOUT: Leveraging Linguistic and Source-based Features for Machine Translation Evaluation
Meritxell Gonzàlez, Alberto Barrón-Cedeño and Lluís Màrquez 394

DiscoTK: Using Discourse Structure for Machine Translation Evaluation
Shafiq Joty, Francisco Guzmán, Lluís Màrquez and Preslav Nakov 402

Tolerant BLEU: a Submission to the WMT14 Metrics Task
Jindřich Libovický and Pavel Pecina 409

BEER: BEtter Evaluation as Ranking
Milos Stanojevic and Khalil Sima'an 414

RED, The DCU-CASICT Submission of Metrics Tasks
Xiaofeng Wu, Hui Yu and Qun Liu 420

Crowdsourcing High-Quality Parallel Data Extraction from Twitter
Wang Ling, Luis Marujo, Chris Dyer, Alan W Black and Isabel Trancoso 426

Using Comparable Corpora to Adapt MT Models to New Domains
Ann Irvine and Chris Callison-Burch 437

Dynamic Topic Adaptation for SMT using Distributional Profiles
Eva Hasler, Barry Haddow and Philipp Koehn 445

Unsupervised Adaptation for Statistical Machine Translation
Saab Mansour and Hermann Ney 457

An Empirical Comparison of Features and Tuning for Phrase-based Machine Translation
Spence Green, Daniel Cer and Christopher Manning 466

Bayesian Reordering Model with Feature Selection
Abdullah Alrajeh and Mahesan Niranjan 477

Augmenting String-to-Tree and Tree-to-String Translation with Non-Syntactic Phrases
Matthias Huck, Hieu Hoang and Philipp Koehn 486

Linear Mixture Models for Robust Machine Translation
Marine Carpuat, Cyril Goutte and George Foster 499

Workshop Program

Thursday, June 26, 2014

9:00–9:10 Opening Remarks

Session 1: Shared Translation Tasks

9:10–9:30 *Efficient Elicitation of Annotations for Human Evaluation of Machine Translation*
Keisuke Sakaguchi, Matt Post and Benjamin Van Durme

9:30–10:00 *Findings of the 2014 Workshop on Statistical Machine Translation*
Ondrej Bojar, Christian Buck, Christian Federmann, Barry Haddow, Philipp Koehn, Johannes Leveling, Christof Monz, Pavel Pecina, Matt Post, Herve Saint-Amand, Radu Soricut, Lucia Specia and Aleš Tamchyna

10:00-10:30 Panel Discussion

10:30–11:00 Coffee

Session 2: Poster Session

11:00-12:30 Shared Task: Translation

Parallel FDA5 for Fast Deployment of Accurate Statistical Machine Translation Systems
Ergun Bicici, Qun Liu and Andy Way

Yandex School of Data Analysis Russian-English Machine Translation System for WMT14
Alexey Borisov and Irina Galinskaya

CimS – The CIS and IMS joint submission to WMT 2014 translating from English into German
Fabienne Cap, Marion Weller, Anita Ramm and Alexander Fraser

English-to-Hindi system description for WMT 2014: Deep Source-Context Features for Moses
Marta R. Costa-jussà, Parth Gupta, Paolo Rosso and Rafael E. Banchs

The KIT-LIMSI Translation System for WMT 2014
Quoc Khanh Do, Teresa Herrmann, Jan Niehues, Alexander Allauzen, François Yvon and Alex Waibel

The IIT Bombay Hindi-English Translation System at WMT 2014
Piyush Dungarwal, Rajen Chatterjee, Abhijit Mishra, Anoop Kunchukuttan, Ritesh Shah and Pushpak Bhattacharyya

Thursday, June 26, 2014 (continued)

Edinburgh's Phrase-based Machine Translation Systems for WMT-14
Nadir Durrani, Barry Haddow, Philipp Koehn and Kenneth Heafield

EU-BRIDGE MT: Combined Machine Translation
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

Phrasal: A Toolkit for New Directions in Statistical Machine Translation
Spence Green, Daniel Cer and Christopher Manning

Anaphora Models and Reordering for Phrase-Based SMT
Christian Hardmeier, Sara Stymne, Jörg Tiedemann, Aaron Smith and Joakim Nivre

The Karlsruhe Institute of Technology Translation Systems for the WMT 2014
Teresa Herrmann, Mohammed Mediani, Eunah Cho, Thanh-Le Ha, Jan Niehues, Isabel Slawik, Yuqi Zhang and Alex Waibel

The DCU-ICTCAS MT system at WMT 2014 on German-English Translation Task
Liangyou Li, Xiaofeng Wu, Santiago Cortes Vaillo, Jun Xie, Andy Way and Qun Liu

The CMU Machine Translation Systems at WMT 2014
Austin Matthews, Waleed Ammar, Archna Bhatia, Weston Feely, Greg Hanneman, Eva Schlinger, Swabha Swayamdipta, Yulia Tsvetkov, Alon Lavie and Chris Dyer

Stanford University's Submissions to the WMT 2014 Translation Task
Julia Neidert, Sebastian Schuster, Spence Green, Kenneth Heafield and Christopher Manning

The RWTH Aachen German-English Machine Translation System for WMT 2014
Stephan Peitz, Joern Wuebker, Markus Freitag and Hermann Ney

Large-scale Exact Decoding: The IMS-TTT submission to WMT14
Daniel Quernheim and Fabienne Cap

Abu-MaTran at WMT 2014 Translation Task: Two-step Data Selection and RBMT-Style Synthetic Rules
Raphael Rubino, Antonio Toral, Víctor M. Sánchez-Cartagena, Jorge Ferrández-Tordera, Sergio Ortiz Rojas, Gema Ramírez-Sánchez, Felipe Sánchez-Martínez and Andy Way

The UA-Prompsit hybrid machine translation system for the 2014 Workshop on Statistical Machine Translation
Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez

Thursday, June 26, 2014 (continued)

Machine Translation and Monolingual Postediting: The AFRL WMT-14 System
Lane Schwartz, Timothy Anderson, Jeremy Gwinnup and Katherine Young

CUNI in WMT14: Chimera Still Awaits Bellerophon
Aleš Tamchyna, Martin Popel, Rudolf Rosa and Ondrej Bojar

Manawi: Using Multi-Word Expressions and Named Entities to Improve Machine Translation
Liling Tan and Santanu Pal

Edinburgh's Syntax-Based Systems at WMT 2014
Philip Williams, Rico Sennrich, Maria Nadejde, Matthias Huck, Eva Hasler and Philipp Koehn

DCU-Lingo24 Participation in WMT 2014 Hindi-English Translation task
Xiaofeng Wu, Rejwanul Haque, Tsuyoshi Okita, Piyush Arora, Andy Way and Qun Liu

11:00-12:30 Shared Task: Medical Translation

Machine Translation of Medical Texts in the Khresmoi Project
Ondřej Dušek, Jan Hajič, Jaroslava Hlaváčová, Michal Novák, Pavel Pecina, Rudolf Rosa, Aleš Tamchyna, Zdeňka Urešová and Daniel Zeman

Postech's System Description for Medical Text Translation Task
Jianri Li, Se-Jong Kim, Hwidong Na and Jong-Hyeok Lee

Domain Adaptation for Medical Text Translation using Web Resources
Yi Lu, Longyue Wang, Derek F. Wong, Lidia S. Chao and Yiming Wang

DCU Terminology Translation System for Medical Query Subtask at WMT14
Tsuyoshi Okita, Ali Vahid, Andy Way and Qun Liu

LIMSI @ WMT'14 Medical Translation Task
Nicolas Pécheux, Li Gong, Quoc Khanh Do, Benjamin Marie, Yulia Ivanishcheva, Alexander Allauzen, Thomas Lavergne, Jan Niehues, Aurélien Max and François Yvon

Combining Domain Adaptation Approaches for Medical Text Translation
Longyue Wang, Yi Lu, Derek F. Wong, Lidia S. Chao, Yiming Wang and Francisco Oliveira

Experiments in Medical Translation Shared Task at WMT 2014
Jian Zhang

Thursday, June 26, 2014 (continued)

12:30–14:00 Lunch

Session 3: Invited Talk

14:00-15:30 *Machine Translation in Academia and in the Commercial World – a Contrastive Perspective.* Alon Lavie, Research Professor – Carnegie Mellon University, Co-founder, President and CTO – Safaba Translation Solutions, Inc.

15:30–16:00 Coffee

Session 4: Evaluation

16:00–16:20 *Randomized Significance Tests in Machine Translation*
Yvette Graham, Nitika Mathur and Timothy Baldwin

16:20–16:40 *Estimating Word Alignment Quality for SMT Reordering Tasks*
Sara Stymne, Jörg Tiedemann and Joakim Nivre

16:40–17:00 *Dependency-based Automatic Enumeration of Semantically Equivalent Word Orders for Evaluating Japanese Translations*
Hideki Isozaki, Natsume Kouchi and Tsutomu Hirao

Friday, June 27, 2014

Session 5: Shared Evaluation Metrics and Quality Estimation Tasks

9:00–9:30 Quality Estimation Shared Task

9:30–9:50 *Results of the WMT14 Metrics Shared Task*
Matous Machacek and Ondrej Bojar

9:50–10:30 Panel

10:30–11:00 Coffee

Friday, June 27, 2014 (continued)

Session 6: Poster Session

11:00–12:30 Shared Task: Quality Estimation

Efforts on Machine Learning over Human-mediated Translation Edit Rate
Eleftherios Avramidis

SHEF-Lite 2.0: Sparse Multi-task Gaussian Processes for Translation Quality Estimation
Daniel Beck, Kashif Shah and Lucia Specia

Referential Translation Machines for Predicting Translation Quality
Ergun Bicici and Andy Way

FBK-UPV-UEdin participation in the WMT14 Quality Estimation shared-task
José Guilherme Camargo de Souza, Jesús González-Rubio, Christian Buck, Marco Turchi and Matteo Negri

Target-Centric Features for Translation Quality Estimation
Chris Hokamp, Iacer Calixto, Joachim Wagner and Jian Zhang

LIG System for Word Level QE task at WMT14
Ngoc Quang Luong, Laurent Besacier and Benjamin Lecouteux

Exploring Consensus in Machine Translation for Quality Estimation
Carolina Scarton and Lucia Specia

LIMSI Submission for WMT'14 QE Task
Guillaume Wisniewski, Nicolas Pécheux, Alexander Allauzen and François Yvon

11:00–12:30 Shared Task: Evaluation Metrics

Parmesan: Meteor without Paraphrases with Paraphrased References
Petra Barancikova

A Systematic Comparison of Smoothing Techniques for Sentence-Level BLEU
Boxing Chen and Colin Cherry

Friday, June 27, 2014 (continued)

VERTa participation in the WMT14 Metrics Task
Elisabet Comelles and Jordi Atserias

Meteor Universal: Language Specific Translation Evaluation for Any Target Language
Michael Denkowski and Alon Lavie

Application of Prize based on Sentence Length in Chunk-based Automatic Evaluation of Machine Translation
Hiroshi Echizen'ya, Kenji Araki and Eduard Hovy

LAYERED: Metric for Machine Translation Evaluation
Shubham Gautam and Pushpak Bhattacharyya

IPA and STOUT: Leveraging Linguistic and Source-based Features for Machine Translation Evaluation
Meritxell Gonzàlez, Alberto Barrón-Cedeño and Lluís Màrquez

DiscoTK: Using Discourse Structure for Machine Translation Evaluation
Shafiq Joty, Francisco Guzmán, Lluís Màrquez and Preslav Nakov

Tolerant BLEU: a Submission to the WMT14 Metrics Task
Jindřich Libovický and Pavel Pecina

BEER: BEtter Evaluation as Ranking
Milos Stanojevic and Khalil Sima'an

RED, The DCU-CASICT Submission of Metrics Tasks
Xiaofeng Wu, Hui Yu and Qun Liu

12:30–14:00 Lunch

Friday, June 27, 2014 (continued)

Session 7: Data and Adaptation

14:00–14:20 *Crowdsourcing High-Quality Parallel Data Extraction from Twitter*
Wang Ling, Luis Marujo, Chris Dyer, Alan W Black and Isabel Trancoso

14:20–14:40 *Using Comparable Corpora to Adapt MT Models to New Domains*
Ann Irvine and Chris Callison-Burch

14:40–15:00 *Dynamic Topic Adaptation for SMT using Distributional Profiles*
Eva Hasler, Barry Haddow and Philipp Koehn

15:00–15:20 *Unsupervised Adaptation for Statistical Machine Translation*
Saab Mansour and Hermann Ney

15:20–16:00 Coffee

Session 8: Translation Models

16:00–16:20 *An Empirical Comparison of Features and Tuning for Phrase-based Machine Translation*
Spence Green, Daniel Cer and Christopher Manning

16:20–16:40 *Bayesian Reordering Model with Feature Selection*
Abdullah Alrajeh and Mahesan Niranjan

16:40–17:00 *Augmenting String-to-Tree and Tree-to-String Translation with Non-Syntactic Phrases*
Matthias Huck, Hieu Hoang and Philipp Koehn

17:00–17:20 *Linear Mixture Models for Robust Machine Translation*
Marine Carpuat, Cyril Goutte and George Foster