

WMT 2011

**Sixth Workshop on
Statistical Machine Translation**

Proceedings of the Workshop

July 30–31, 2011

With support from EUROMATRIXPLUS,
an EU Framework Programme 7 funded project.

©2011 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 978-1-937284-12-1/1-937284-12-3

Introduction

The EMNLP 2011 Workshop on Statistical Machine Translation (WMT-2011) took place on Saturday and Sunday, July 30–31 in Edinburgh, Scotland, immediately following the Conference on Empirical Methods on Natural Language Processing (EMNLP) 2011, which was hosted by the University of Edinburgh.

This is the seventh time this workshop has been held. The first time was in 2005 as part of the ACL 2005 Workshop on Building and Using Parallel Texts. In the following years the Workshop on Statistical Machine Translation was held at HLT-NAACL 2006 in New York City, US, ACL 2007 in Prague, Czech Republic, ACL 2008, Columbus, Ohio, US, EACL 2009 in Athens, Greece, and ACL 2010 in Uppsala, Sweden.

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 a shared task that brought together machine translation systems for an evaluation on previously unseen data. The results of the shared task were announced at the workshop, and these proceedings also include an overview paper for the shared task 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 some 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 This year we have received 42 full paper submissions (not counting withdrawn submissions) and 47 shared task submissions. In total WMT-2011 featured 18 full paper oral presentations and 47 shared task poster presentations.

The invited talk was given by William Lewis (Microsoft Research), Robert Munro (Stanford University), and Stephan Vogel (Carnegie Mellon University).

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.

Chris Callison-Burch, Philipp Koehn, Christof Monz, and Omar F. Zaidan

WMT 5-year Retrospective Best Paper Award

Since this is the Sixth Workshop on Statistical Machine Translation, we have decided to create a WMT 5-year Retrospective Best Paper Award, to be given to the best paper that was published at the first Workshop on Statistical Machine Translation, which was held at HLT-NAACL 2006 in New York. The goals of this retrospective award are to recognize high-quality work that has stood the test of time, and to highlight the excellent work that appears at WMT.

The WMT11 program committee voted on the best paper from a list of six nominated papers. Five of these were nominated by high citation counts, which we defined as having 10 or more citations in the ACL anthology network (excluding self-citations), and more than 30 citations on Google Scholar. We also opened the nomination process to the committee, which yielded one further nomination for a paper that did not reach the citation threshold but was deemed to be excellent.

The program committee decided to award the WMT 5-year Retrospective Best Paper Award to:

Andreas Zollmann and Ashish Venugopal. 2006. *Syntax Augmented Machine Translation via Chart Parsing*. In Proceedings of the Workshop on Statistical Machine Translation. Pages 138–141.

This short paper described Zollmann and Venugopal’s entry into the WMT06 shared translation task. Their system introduced a parsing-based machine translation system. Like David Chiang’s Hiero system, Zollmann and Venugopal’s system used synchronous context free grammars (SCFGs). Instead of using a single non-terminal symbol, X, Zollmann and Venugopal’s SCFG rules contained linguistically informed non-terminal symbols that were extracted from a parsed parallel corpus.

This paper was one of the first publications to demonstrate that syntactically-informed approaches to statistical machine translation could achieve translation quality that was comparable to – or even better than – state-of-the-art phrase-based and hierarchical phrase-based approaches to machine translation. Zollmann and Venugopal’s approach has influenced a number of researchers, and has been integrated into open source translation software like the Joshua and Moses decoders.

In many ways this paper represents the ideals of the WMT workshops. It introduced a novel approach to machine translation and demonstrated its value empirically by comparing it to other state-of-the-art systems on a public data set.

Congratulations to Andreas Zollmann and Ashish Venugopal for their excellent work!

Organizers:

Chris Callison-Burch (Johns Hopkins University)
Philipp Koehn (University of Edinburgh)
Christof Monz (University of Amsterdam)
Omar F. Zaidan (Johns Hopkins University)

Invited Talk:

William Lewis (Microsoft Research)
Robert Munro (Stanford University)
Stephan Vogel (Carnegie Mellon University)

Program Committee:

Lars Ahrenberg (Linköping University)
Nicola Bertoldi (FBK)
Graeme Blackwood (University of Cambridge)
Michael Bloodgood (University of Maryland)
Ondrej Bojar (Charles University)
Thorsten Brants (Google)
Chris Brockett (Microsoft)
Nicola Cancedda (Xerox)
Marine Carpuat (National Research Council Canada)
Simon Carter (University of Amsterdam)
Francisco Casacuberta (University of Valencia)
Daniel Cer (Stanford University)
Mauro Cettolo (FBK)
Boxing Chen (National Research Council Canada)
Colin Cherry (National Research Council Canada)
David Chiang (ISI)
Jon Clark (Carnegie Mellon University)
Stephen Clark (University of Cambridge)
Christophe Costa Florencio (University of Amsterdam)
Michael Denkowski (Carnegie Mellon University)
Kevin Duh (NTT)
Chris Dyer (Carnegie Mellon University)
Marc Dymetman (Xerox)
Marcello Federico (FBK)

Andrew Finch (NICT)
Jose Fonollosa (University of Catalonia)
George Foster (National Research Council Canada)
Alex Fraser (University of Stuttgart)
Michel Galley (Microsoft)
Niyu Ge (IBM)
Dmitriy Genzel (Google)
Ulrich Germann (University of Toronto)
Kevin Gimpel (Carnegie Mellon University)
Adria de Gispert (University of Cambridge)
Spence Green (Stanford University)
Nizar Habash (Columbia University)
Keith Hall (Google)
Greg Hanneman (Carnegie Mellon University)
Kenneth Heafield (Carnegie Mellon University)
John Henderson (MITRE)
Howard Johnson (National Research Council Canada)
Doug Jones (Lincoln Labs MIT)
Damianos Karakos (Johns Hopkins University)
Maxim Khalilov (University of Amsterdam)
Roland Kuhn (National Research Council Canada)
Shankar Kumar (Google)
Philippe Langlais (Univeristy of Montreal)
Adam Lopez (Johns Hopkins University)
Wolfgang Macherey (Google)
Nitin Madnani (Educational Testing Service)
Daniel Marcu (Language Weaver)
Yuval Marton (IBM)
Lambert Mathias (Nuance)
Spyros Matsoukas (Raytheon BBN Technologies)
Arne Mauser (RWTH Aachen)
Arul Menezes (Microsoft)
Bob Moore (Google)
Smaranda Muresan (Rutgers University)
Kemal Oflazer (Carnegie Mellon University)
Miles Osborne (University of Edinburgh)
Matt Post (Johns Hopkins University)
Chris Quirk (Microsoft)
Antti-Veikko Rosti (Raytheon BBN Technologies)
Salim Roukos (IBM)
Anoop Sarkar (Simon Fraser University)
Holger Schwenk (University of Le Mans)

Jean Senellart (Systran)
Khalil Simaan (University of Amsterdam)
Michel Simard (National Research Council Canada)
David Smith (University of Massachusetts Amherst)
Matthew Snover (Raytheon BBN Technologies)
Joerg Tiedemann (Uppsala University)
Christoph Tillmann (IBM)
Dan Tufis (Romanian Academy)
Jakob Uszkoreit (Google)
Masao Utiyama (NICT)
David Vilar (RWTH Aachen)
Clare Voss (Army Research Labs)
Taro Watanabe (NTT)
Andy Way (Dublin City University)
Jinxi Xu (Raytheon BBN Technologies)
Sirvan Yahyaei (Queen Mary, University of London)
Daniel Zeman (Charles University)
Richard Zens (Google)
Bing Zhang (Raytheon BBN Technologies)

Table of Contents

<i>A Grain of Salt for the WMT Manual Evaluation</i>	
Ondřej Bojar, Miloš Ercegovčević, Martin Popel and Omar Zaidan	1
<i>A Lightweight Evaluation Framework for Machine Translation Reordering</i>	
David Talbot, Hideto Kazawa, Hiroshi Ichikawa, Jason Katz-Brown, Masakazu Seno and Franz Och	12
<i>Findings of the 2011 Workshop on Statistical Machine Translation</i>	
Chris Callison-Burch, Philipp Koehn, Christof Monz and Omar Zaidan	22
<i>Evaluate with Confidence Estimation: Machine ranking of translation outputs using grammatical features</i>	
Eleftherios Avramidis, Maja Popović, David Vilar and Aljoscha Burchardt	65
<i>AMBER: A Modified BLEU, Enhanced Ranking Metric</i>	
Boxing Chen and Roland Kuhn	71
<i>TESLA at WMT 2011: Translation Evaluation and Tunable Metric</i>	
Daniel Dahlmeier, Chang Liu and Hwee Tou Ng	78
<i>Meteor 1.3: Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems</i>	
Michael Denkowski and Alon Lavie	85
<i>Approximating a Deep-Syntactic Metric for MT Evaluation and Tuning</i>	
Matouš Macháček and Ondřej Bojar	92
<i>Evaluation without references: IBM1 scores as evaluation metrics</i>	
Maja Popović, David Vilar, Eleftherios Avramidis and Aljoscha Burchardt	99
<i>Morphemes and POS tags for n-gram based evaluation metrics</i>	
Maja Popović	104
<i>E-rating Machine Translation</i>	
Kristen Parton, Joel Tetreault, Nitin Madnani and Martin Chodorow	108
<i>TINE: A Metric to Assess MT Adequacy</i>	
Miguel Rios, Wilker Aziz and Lucia Specia	116
<i>Regression and Ranking based Optimisation for Sentence Level MT Evaluation</i>	
Xingyi Song and Trevor Cohn	123
<i>MAISE: A Flexible, Configurable, Extensible Open Source Package for Mass AI System Evaluation</i>	
Omar Zaidan	130

<i>MANY improvements for WMT'11</i>	
Loïc Barrault	135
<i>The UPV-PRHLT combination system for WMT 2011</i>	
Jesús González-Rubio and Francisco Casacuberta	140
<i>CMU System Combination in WMT 2011</i>	
Kenneth Heafield and Alon Lavie	145
<i>The RWTH System Combination System for WMT 2011</i>	
Gregor Leusch, Markus Freitag and Hermann Ney	152
<i>Expected BLEU Training for Graphs: BBN System Description for WMT11 System Combination Task</i>	
Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz	159
<i>The UZH System Combination System for WMT 2011</i>	
Rico Sennrich	166
<i>Description of the JHU System Combination Scheme for WMT 2011</i>	
Daguang Xu, Yuan Cao and Damianos Karakos	171
<i>Multiple-stream Language Models for Statistical Machine Translation</i>	
Abby Levenberg, Miles Osborne and David Matthews	177
<i>KenLM: Faster and Smaller Language Model Queries</i>	
Kenneth Heafield	187
<i>Wider Context by Using Bilingual Language Models in Machine Translation</i>	
Jan Niehues, Teresa Herrmann, Stephan Vogel and Alex Waibel	198
<i>A Minimally Supervised Approach for Detecting and Ranking Document Translation Pairs</i>	
Kriste Krstovski and David A. Smith	207
<i>Agreement Constraints for Statistical Machine Translation into German</i>	
Philip Williams and Philipp Koehn	217
<i>Fuzzy Syntactic Reordering for Phrase-based Statistical Machine Translation</i>	
Jacob Andreas, Nizar Habash and Owen Rambow	227
<i>Filtering Antonymous, Trend-Contrasting, and Polarity-Dissimilar Distributional Paraphrases for Improving Statistical Machine Translation</i>	
Yuval Marton, Ahmed El Kholy and Nizar Habash	237
<i>Productive Generation of Compound Words in Statistical Machine Translation</i>	
Sara Stymne and Nicola Cancedda	250
<i>SampleRank Training for Phrase-Based Machine Translation</i>	
Barry Haddow, Abhishek Arun and Philipp Koehn	261

<i>Instance Selection for Machine Translation using Feature Decay Algorithms</i> Ergun Biciçi and Deniz Yuret	272
<i>Investigations on Translation Model Adaptation Using Monolingual Data</i> Patrik Lambert, Holger Schwenk, Christophe Servan and Sadaf Abdul-Rauf	284
<i>Topic Adaptation for Lecture Translation through Bilingual Latent Semantic Models</i> Nick Ruiz and Marcello Federico	294
<i>Personal Translator at WMT2011</i> Vera Aleksic and Gregor Thurmair	303
<i>LIMSI @ WMT11</i> Alexandre Allauzen, H�el�ene Bonneau-Maynard, Hai-Son Le, Aur�elien Max, Guillaume Wisniewski, Fran�ois Yvon, Gilles Adda, Josep Maria Crego, Adrien Lardilleux, Thomas Lavergne and Artem Sokolov	309
<i>Shallow Semantic Trees for SMT</i> Wilker Aziz, Miguel Rios and Lucia Specia	316
<i>RegMT System for Machine Translation, System Combination, and Evaluation</i> Ergun Biciçi and Deniz Yuret	323
<i>Improving Translation Model by Monolingual Data</i> Ondr�ej Bojar and Aleř Tamchyna	330
<i>The CMU-ARK German-English Translation System</i> Chris Dyer, Kevin Gimpel, Jonathan H. Clark and Noah A. Smith	337
<i>Noisy SMS Machine Translation in Low-Density Languages</i> Vladimir Eidelman, Kristy Hollingshead and Philip Resnik	344
<i>Stochastic Parse Tree Selection for an Existing RBMT System</i> Christian Federmann and Sabine Hunsicker	351
<i>Joint WMT Submission of the QUAERO Project</i> Markus Freitag, Gregor Leusch, Joern Wuebker, Stephan Peitz, Hermann Ney, Teresa Herrmann, Jan Niehues, Alex Waibel, Alexandre Allauzen, Gilles Adda, Josep Maria Crego, Bianca Buschbeck, Tonio Wandmacher and Jean Senellart	358
<i>CMU Syntax-Based Machine Translation at WMT 2011</i> Greg Hanneman and Alon Lavie	365
<i>The Uppsala-FBK systems at WMT 2011</i> Christian Hardmeier, J�org Tiedemann, Markus Saers, Marcello Federico and Mathur Prashant	372
<i>The Karlsruhe Institute of Technology Translation Systems for the WMT 2011</i> Teresa Herrmann, Mohammed Mediani, Jan Niehues and Alex Waibel	379

<i>CMU Haitian Creole-English Translation System for WMT 2011</i>	
Sanjika Hewavitharana, Nguyen Bach, Qin Gao, Vamshi Ambati and Stephan Vogel	386
<i>Experiments with word alignment, normalization and clause reordering for SMT between English and German</i>	
Maria Holmqvist, Sara Stymne and Lars Ahrenberg	393
<i>The Value of Monolingual Crowdsourcing in a Real-World Translation Scenario: Simulation using Haitian Creole Emergency SMS Messages</i>	
Chang Hu, Philip Resnik, Yakov Kronrod, Vladimir Eidelman, Olivia Buzek and Benjamin B. Bederson	399
<i>The RWTH Aachen Machine Translation System for WMT 2011</i>	
Matthias Huck, Joern Wuebker, Christoph Schmidt, Markus Freitag, Stephan Peitz, Daniel Stein, Arnaud Dagnelies, Saab Mansour, Gregor Leusch and Hermann Ney	405
<i>ILLC-UvA translation system for EMNLP-WMT 2011</i>	
Maxim Khalilov and Khalil Sima'an	413
<i>UPM system for the translation task</i>	
Verónica López-Ludeña and Rubén San-Segundo	420
<i>Two-step translation with grammatical post-processing</i>	
David Mareček, Rudolf Rosa, Petra Galuščáková and Ondřej Bojar	426
<i>Influence of Parser Choice on Dependency-Based MT</i>	
Martin Popel, David Mareček, Nathan Green and Zdenek Zabokrtsky	433
<i>The LIGA (LIG/LIA) Machine Translation System for WMT 2011</i>	
Marion Potet, Raphaël Rubino, Benjamin Lecouteux, Stéphane Huet, Laurent Besacier, Hervé Blanchon and Fabrice Lefèvre	440
<i>Factored Translation with Unsupervised Word Clusters</i>	
Christian Rishøj and Anders Søgaard	447
<i>The BM-I2R Haitian-Créole-to-English translation system description for the WMT 2011 evaluation campaign</i>	
Marta R. Costa-jussà and Rafael E. Banchs	452
<i>The Universitat d'Alacant hybrid machine translation system for WMT 2011</i>	
Víctor M. Sánchez-Cartagena, Felipe Sánchez-Martínez and Juan Antonio Pérez-Ortiz	457
<i>LIUM's SMT Machine Translation Systems for WMT 2011</i>	
Holger Schwenk, Patrik Lambert, Loïc Barrault, Christophe Servan, Sadaf Abdul-Rauf, Haithem Afli and Kashif Shah	464
<i>Spell Checking Techniques for Replacement of Unknown Words and Data Cleaning for Haitian Creole SMS Translation</i>	
Sara Stymne	470

<i>Joshua 3.0: Syntax-based Machine Translation with the Thrax Grammar Extractor</i>	
Jonathan Weese, Juri Ganitkevitch, Chris Callison-Burch, Matt Post and Adam Lopez	478
<i>DFKI Hybrid Machine Translation System for WMT 2011 - On the Integration of SMT and RBMT</i>	
Jia Xu, Hans Uszkoreit, Casey Kennington, David Vilar and Xiaojun Zhang	485
<i>CEU-UPV English-Spanish system for WMT11</i>	
Francisco Zamora-Martinez and Maria Jose Castro-Bleda	490
<i>Hierarchical Phrase-Based MT at the Charles University for the WMT 2011 Shared Task</i>	
Daniel Zeman	496
<i>Crisis MT: Developing A Cookbook for MT in Crisis Situations</i>	
William Lewis, Robert Munro and Stephan Vogel	501
<i>Generative Models of Monolingual and Bilingual Gappy Patterns</i>	
Kevin Gimpel and Noah A. Smith	512
<i>Extraction Programs: A Unified Approach to Translation Rule Extraction</i>	
Mark Hopkins, Greg Langmead and Tai Vo	523
<i>Bayesian Extraction of Minimal SCFG Rules for Hierarchical Phrase-based Translation</i>	
Baskaran Sankaran, Gholamreza Haffari and Anoop Sarkar	533
<i>From n-gram-based to CRF-based Translation Models</i>	
Thomas Lavergne, Alexandre Allauzen, Josep Maria Crego and François Yvon	542

Conference Program

Saturday, July 30, 2011

Session 1: Shared Tasks and Evaluation

- 9:00–9:20 *A Grain of Salt for the WMT Manual Evaluation*
Ondřej Bojar, Miloš Ercegovič, Martin Popel and Omar Zaidan
- 9:20–9:40 *A Lightweight Evaluation Framework for Machine Translation Reordering*
David Talbot, Hideto Kazawa, Hiroshi Ichikawa, Jason Katz-Brown, Masakazu Seno and Franz Och
- 9:40–10:30 *Findings of the 2011 Workshop on Statistical Machine Translation*
Chris Callison-Burch, Philipp Koehn, Christof Monz and Omar Zaidan
- 10:30–11:00 Coffee

Session 2: Shared Metrics and System Combination Tasks

- 11:00–12:40 Poster Session Metrics
- Evaluate with Confidence Estimation: Machine ranking of translation outputs using grammatical features*
Eleftherios Avramidis, Maja Popović, David Vilar and Aljoscha Burchardt
- AMBER: A Modified BLEU, Enhanced Ranking Metric*
Boxing Chen and Roland Kuhn
- TESLA at WMT 2011: Translation Evaluation and Tunable Metric*
Daniel Dahlmeier, Chang Liu and Hwee Tou Ng
- Meteor 1.3: Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems*
Michael Denkowski and Alon Lavie
- Approximating a Deep-Syntactic Metric for MT Evaluation and Tuning*
Matouš Macháček and Ondřej Bojar
- Evaluation without references: IBM1 scores as evaluation metrics*
Maja Popović, David Vilar, Eleftherios Avramidis and Aljoscha Burchardt

Saturday, July 30, 2011 (continued)

Morphemes and POS tags for n-gram based evaluation metrics

Maja Popović

E-rating Machine Translation

Kristen Parton, Joel Tetreault, Nitin Madnani and Martin Chodorow

TINE: A Metric to Assess MT Adequacy

Miguel Rios, Wilker Aziz and Lucia Specia

Regression and Ranking based Optimisation for Sentence Level MT Evaluation

Xingyi Song and Trevor Cohn

MAISE: A Flexible, Configurable, Extensible Open Source Package for Mass AI System Evaluation

Omar Zaidan

11:00–12:40 Poster Session System Combination

MANY improvements for WMT'11

Loïc Barrault

The UPV-PRHLT combination system for WMT 2011

Jesús González-Rubio and Francisco Casacuberta

CMU System Combination in WMT 2011

Kenneth Heafield and Alon Lavie

The RWTH System Combination System for WMT 2011

Gregor Leusch, Markus Freitag and Hermann Ney

Expected BLEU Training for Graphs: BBN System Description for WMT11 System Combination Task

Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz

The UZH System Combination System for WMT 2011

Rico Sennrich

Description of the JHU System Combination Scheme for WMT 2011

Daguang Xu, Yuan Cao and Damianos Karakos

Saturday, July 30, 2011 (continued)

12:40–14:00 Lunch

Session 3: Language Models and Document Alignment

14:00–14:25 *Multiple-stream Language Models for Statistical Machine Translation*
Abby Levenberg, Miles Osborne and David Matthews

14:25–14:50 *KenLM: Faster and Smaller Language Model Queries*
Kenneth Heafield

14:50–15:15 *Wider Context by Using Bilingual Language Models in Machine Translation*
Jan Niehues, Teresa Herrmann, Stephan Vogel and Alex Waibel

15:15–15:40 *A Minimally Supervised Approach for Detecting and Ranking Document Translation Pairs*
Kriste Krstovski and David A. Smith

15:40–16:10 Coffee

Session 4: Syntax and Semantics

16:10–16:35 *Agreement Constraints for Statistical Machine Translation into German*
Philip Williams and Philipp Koehn

16:35–17:00 *Fuzzy Syntactic Reordering for Phrase-based Statistical Machine Translation*
Jacob Andreas, Nizar Habash and Owen Rambow

17:00–17:25 *Filtering Antonymous, Trend-Contrasting, and Polarity-Dissimilar Distributional Paraphrases for Improving Statistical Machine Translation*
Yuval Marton, Ahmed El Kholy and Nizar Habash

17:25–17:50 *Productive Generation of Compound Words in Statistical Machine Translation*
Sara Stymne and Nicola Cancedda

Sunday, July 31, 2011

Session 5: Machine Learning and Adaptation

- 9:00–10:25 *SampleRank Training for Phrase-Based Machine Translation*
Barry Haddow, Abhishek Arun and Philipp Koehn
- 9:25–10:50 *Instance Selection for Machine Translation using Feature Decay Algorithms*
Ergun Bici and Deniz Yuret
- 9:50–10:15 *Investigations on Translation Model Adaptation Using Monolingual Data*
Patrik Lambert, Holger Schwenk, Christophe Servan and Sadaf Abdul-Rauf
- 10:15–10:40 *Topic Adaptation for Lecture Translation through Bilingual Latent Semantic Models*
Nick Ruiz and Marcello Federico
- 10:40–11:00 Coffee

Session 6: Shared Translation Task

- 11:00–12:40 Poster Presentations

Personal Translator at WMT2011
Vera Aleksic and Gregor Thurmair

LIMSI @ WMT11
Alexandre Allauzen, H el ene Bonneau-Maynard, Hai-Son Le, Aur elien Max, Guillaume Wisniewski, Fran ois Yvon, Gilles Adda, Josep Maria Crego, Adrien Lardilleux, Thomas Lavergne and Artem Sokolov

Shallow Semantic Trees for SMT
Wilker Aziz, Miguel Rios and Lucia Specia

RegMT System for Machine Translation, System Combination, and Evaluation
Ergun Bici and Deniz Yuret

Improving Translation Model by Monolingual Data
Ondr ej Bojar and Aleř Tamchyna

Sunday, July 31, 2011 (continued)

The CMU-ARK German-English Translation System

Chris Dyer, Kevin Gimpel, Jonathan H. Clark and Noah A. Smith

Noisy SMS Machine Translation in Low-Density Languages

Vladimir Eidelman, Kristy Hollingshead and Philip Resnik

Stochastic Parse Tree Selection for an Existing RBMT System

Christian Federmann and Sabine Hunsicker

Joint WMT Submission of the QUAERO Project

Markus Freitag, Gregor Leusch, Joern Wuebker, Stephan Peitz, Hermann Ney, Teresa Herrmann, Jan Niehues, Alex Waibel, Alexandre Allauzen, Gilles Adda, Josep Maria Crego, Bianka Buschbeck, Tonio Wandmacher and Jean Senellart

CMU Syntax-Based Machine Translation at WMT 2011

Greg Hanneman and Alon Lavie

The Uppsala-FBK systems at WMT 2011

Christian Hardmeier, Jörg Tiedemann, Markus Saers, Marcello Federico and Mathur Prashant

The Karlsruhe Institute of Technology Translation Systems for the WMT 2011

Teresa Herrmann, Mohammed Mediani, Jan Niehues and Alex Waibel

CMU Haitian Creole-English Translation System for WMT 2011

Sanjika Hewavitharana, Nguyen Bach, Qin Gao, Vamshi Ambati and Stephan Vogel

Experiments with word alignment, normalization and clause reordering for SMT between English and German

Maria Holmqvist, Sara Stymne and Lars Ahrenberg

The Value of Monolingual Crowdsourcing in a Real-World Translation Scenario: Simulation using Haitian Creole Emergency SMS Messages

Chang Hu, Philip Resnik, Yakov Kronrod, Vladimir Eidelman, Olivia Buzek and Benjamin B. Bederson

The RWTH Aachen Machine Translation System for WMT 2011

Matthias Huck, Joern Wuebker, Christoph Schmidt, Markus Freitag, Stephan Peitz, Daniel Stein, Arnaud Dagnelies, Saab Mansour, Gregor Leusch and Hermann Ney

ILLC-UvA translation system for EMNLP-WMT 2011

Maxim Khalilov and Khalil Sima'an

Sunday, July 31, 2011 (continued)

UPM system for the translation task

Verónica López-Ludeña and Rubén San-Segundo

Two-step translation with grammatical post-processing

David Mareček, Rudolf Rosa, Petra Galuščáková and Ondřej Bojar

Influence of Parser Choice on Dependency-Based MT

Martin Popel, David Mareček, Nathan Green and Zdenek Zabokrtsky

The LIGA (LIG/LIA) Machine Translation System for WMT 2011

Marion Potet, Raphaël Rubino, Benjamin Lecouteux, Stéphane Huet, Laurent Besacier, Hervé Blanchon and Fabrice Lefèvre

Factored Translation with Unsupervised Word Clusters

Christian Rishøj and Anders Søgaard

The BM-I2R Haitian-Créole-to-English translation system description for the WMT 2011 evaluation campaign

Marta R. Costa-jussà and Rafael E. Banchs

The Universitat d'Alacant hybrid machine translation system for WMT 2011

Víctor M. Sánchez-Cartagena, Felipe Sánchez-Martínez and Juan Antonio Pérez-Ortiz

LIUM's SMT Machine Translation Systems for WMT 2011

Holger Schwenk, Patrik Lambert, Loïc Barrault, Christophe Servan, Sadaf Abdul-Rauf, Haithem Afli and Kashif Shah

Spell Checking Techniques for Replacement of Unknown Words and Data Cleaning for Haitian Creole SMS Translation

Sara Stymne

Joshua 3.0: Syntax-based Machine Translation with the Thrax Grammar Extractor

Jonathan Weese, Juri Ganitkevitch, Chris Callison-Burch, Matt Post and Adam Lopez

DFKI Hybrid Machine Translation System for WMT 2011 - On the Integration of SMT and RBMT

Jia Xu, Hans Uszkoreit, Casey Kennington, David Vilar and Xiaojun Zhang

CEU-UPV English-Spanish system for WMT11

Francisco Zamora-Martinez and Maria Jose Castro-Bleda

Sunday, July 31, 2011 (continued)

Hierarchical Phrase-Based MT at the Charles University for the WMT 2011 Shared Task
Daniel Zeman

12:40–14:00 Lunch

Invited Talk

14:00–15:40 *Crisis MT: Developing A Cookbook for MT in Crisis Situations*
William Lewis, Robert Munro and Stephan Vogel

15:40–16:10 Coffee

Session 7: Translation Models

16:10–16:35 *Generative Models of Monolingual and Bilingual Gappy Patterns*
Kevin Gimpel and Noah A. Smith

16:35–17:00 *Extraction Programs: A Unified Approach to Translation Rule Extraction*
Mark Hopkins, Greg Langmead and Tai Vo

17:00–17:25 *Bayesian Extraction of Minimal SCFG Rules for Hierarchical Phrase-based Translation*
Baskaran Sankaran, Gholamreza Haffari and Anoop Sarkar

17:25–17:50 *From n-gram-based to CRF-based Translation Models*
Thomas Lavergne, Alexandre Allauzen, Josep Maria Crego and François Yvon

