3rd Workshop on Statistical Machine Translation 2008

(WMT 08)

Held in conjunction with ACL-HLT 2008

Columbus, Ohio, USA
19 June 2008

ISBN: 978-1-61738-550-6

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© (2008) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2010)

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

An Empirical Study in Source Word Deletion for Phrase-Based Statistical Machine Translation
Chi-Ho Li, Hailei Zhang, Dongdong Zhang, Mu Li and Ming Zhou 1

Rich Source-Side Context for Statistical Machine Translation
Kevin Gimpel and Noah A. Smith 9

Discriminative Word Alignment via Alignment Matrix Modeling
Jan Niehues and Stephan Vogel 18

Regularization and Search for Minimum Error Rate Training
Daniel Cer, Dan Jurafsky and Christopher D. Manning 26

Learning Performance of a Machine Translation System: a Statistical and Computational Analysis
Marco Turchi, Tijl De Bie and Nello Cristianini 35

Using Syntax to Improve Word Alignment Precision for Syntax-Based Machine Translation
Victoria Fossum, Kevin Knight and Steven Abney 44

Using Shallow Syntax Information to Improve Word Alignment and Reordering for SMT
Josep M. Crego and Nizar Habash 53

Improved Tree-to-String Transducer for Machine Translation
Ding Liu and Daniel Gildea 62

Further Meta-Evaluation of Machine Translation
Chris Callison-Burch, Cameron Fordyce, Philipp Koehn, Christof Monz and Josh Schroeder . . 70

Limsi's Statistical Translation Systems for WMT'08
Daniel Déchelotte, Gilles Adda, Alexandre Allauzen, Hélène Bonneau-Maynard, Olivier Galibert, Jean-Luc Gauvain, Philippe Langlais and François Yvon 107

The MetaMorpho Translation System
Attila Novák, László Tihanyi and Gábor Prószéky 111

Meteor, M-BLEU and M-TER: Evaluation Metrics for High-Correlation with Human Rankings of Machine Translation Output
Abhaya Agarwal and Alon Lavie 115

First Steps towards a General Purpose French/English Statistical Machine Translation System
Holger Schwenk, Jean-Baptiste Fouet and Jean Senellart 119

The University of Washington Machine Translation System for ACL WMT 2008
Amittai Axelrod, Mei Yang, Kevin Duh and Katrin Kirchhoff 123

The TALP-UPC Ngram-Based Statistical Machine Translation System for ACL-WMT 2008
Maxim Khalilov, Adolfo Hernández H., Marta R. Costa-jussà, Josep M. Crego, Carlos A. Henríquez Q., Patrik Lambert, José A. R. Fonollosa, José B. Mariño and Rafael E. Banchs 127

European Language Translation with Weighted Finite State Transducers: The CUED MT System for the 2008 ACL Workshop on SMT
Graeme Blackwood, Adrià de Gispert, Jamie Brunning and William Byrne 131

Effects of Morphological Analysis in Translation between German and English
Sara Stymne, Maria Holmqvist and Lars Ahrenberg 135

Towards better Machine Translation Quality for the German-English Language Pairs
Philipp Koehn, Abhishek Arun and Hieu Hoang 139

Phrase-Based and Deep Syntactic English-to-Czech Statistical Machine Translation
Ondřej Bojar and Jan Hajič 143

Improving English-Spanish Statistical Machine Translation: Experiments in Domain Adaptation, Sentence Paraphrasing, Tokenization, and Recasing
Preslav Nakov 147

Improving Word Alignment with Language Model Based Confidence Scores
Nguyen Bach, Qin Gao and Stephan Vogel 151

Kernel Regression Framework for Machine Translation: UCL System Description for WMT 2008 Shared Translation Task
Zhuoran Wang and John Shawe-Taylor 155

Using Syntactic Coupling Features for Discriminating Phrase-Based Translations (WMT-08 Shared Translation Task)
Vassilina Nikoulina and Marc Dymetman 159

Statistical Transfer Systems for French-English and German-English Machine Translation
Greg Hanneman, Edmund Huber, Abhaya Agarwal, Vamshi Ambati, Alok Parlikar, Erik Peterson and Alon Lavie 163

TectoMT: Highly Modular MT System with Tectogrammatics Used as Transfer Layer
Zdenek Zabokrtsky, Jan Ptacek and Petr Pajas 167

MaTrEx: The DCU MT System for WMT 2008
John Tinsley, Yanjun Ma, Sylwia Ozdowska and Andy Way 171

Can we Relearn an RBMT System?
Loïc Dugast, Jean Senellart and Philipp Koehn 175

Using Moses to Integrate Multiple Rule-Based Machine Translation Engines into a Hybrid System
Andreas Eisele, Christian Federmann, Hervé Saint-Amand, Michael Jellinghaus, Teresa Herrmann and Yu Chen 179

Incremental Hypothesis Alignment for Building Confusion Networks with Application to Machine Translation System Combination
Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz 183

The Role of Pseudo References in MT Evaluation
Joshua Albrecht and Rebecca Hwa 187

Ranking vs. Regression in Machine Translation Evaluation
Kevin Duh 191

A Smorgasbord of Features for Automatic MT Evaluation
Jesús Giménez and Lluís Màrquez 195

Fast, Easy, and Cheap: Construction of Statistical Machine Translation Models with MapReduce
Chris Dyer, Aaron Cordova, Alex Mont and Jimmy Lin 199

Dynamic Model Interpolation for Statistical Machine Translation
Andrew Finch and Eiichiro Sumita 208

Improved Statistical Machine Translation by Multiple Chinese Word Segmentation
Ruiqiang Zhang, Keiji Yasuda and Eiichiro Sumita 216

Optimizing Chinese Word Segmentation for Machine Translation Performance
Pi-Chuan Chang, Michel Galley and Christopher D. Manning 224