8th Workshop on Syntax, Semantics and Structure in Statistical Translation 2014

(SSST-8)

Held at the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)

Doha, Qatar
25 October 2014

Editors:

D. Wu
M. Carpuat
X. Carreras
E. M. Vecchi

ISBN: 978-1-63439-482-6

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

Vector Space Models for Phrase-based Machine Translation
Tamer Alkhouli, Andreas Guta and Hermann Ney 1

Bilingual Markov Reordering Labels for Hierarchical SMT
Gideon Maillette de Buy Wenniger and Khalil Sima'an 11

Better Semantic Frame Based MT Evaluation via Inversion Transduction Grammars
Dekai Wu, Chi-kiu Lo, Meriem Beloucif and Markus Saers 22

Rule-based Syntactic Preprocessing for Syntax-based Machine Translation
Yuto Hatakoshi, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura 34

Applying HMEANT to English-Russian Translations
Alexander Chuchunkov, Alexander Tarelkin and Irina Galinskaya 43

Reducing the Impact of Data Sparsity in Statistical Machine Translation
Karan Singla, Kunal Sachdeva, Srinivas Bangalore, Dipti Misra Sharma and Diksha Yadav 51

Expanding the Language model in a low-resource hybrid MT system
George Tambouratzis, Sokratis Sofianopoulos and Marina Vassiliou 57

Syntax and Semantics in Quality Estimation of Machine Translation
Rasoul Kaljahi, Jennifer Foster and Johann Roturier 67

Overcoming the Curse of Sentence Length for Neural Machine Translation using Automatic Segmentation
Jean Pouget-Abadie, Dzmitry Bahdanau, Bart van Merrienboer, Kyunghyun Cho and Yoshua Bengio 78

Ternary Segmentation for Improving Search in Top-down Induction of Segmental ITGs
Markus Saers and Dekai Wu 86

A CYK+ Variant for SCFG Decoding Without a Dot Chart
Rico Sennrich 94

On the Properties of Neural Machine Translation: Encoder–Decoder Approaches
Kyunghyun Cho, Bart van Merrienboer, Dzmitry Bahdanau and Yoshua Bengio 103

Transduction Recursive Auto-Associative Memory: Learning Bilingual Compositional Distributed Vector Representations of Inversion Transduction Grammars
Karteek Addanki and Dekai Wu 112

Transformation and Decomposition for Efficiently Implementing and Improving Dependency-to-String Model In Moses
Liangyou Li, Jun Xie, Andy Way and Qun Liu 122

Word's Vector Representations meet Machine Translation
Eva Martinez Garcia, Jörg Tiedemann, Cristina España-Bonet and Lluís Màrquez 132

Context Sense Clustering for Translation
João Casteleiro, Gabriel Lopes and Joaquim Silva 135

Evaluating Word Order Recursively over Permutation-Forests
Miloš Stanojević and Khalil Sima'an 138

Preference Grammars and Soft Syntactic Constraints for GHKM Syntax-based Statistical Machine Translation
Matthias Huck, Hieu Hoang and Philipp Koehn 148

How Synchronous are Adjuncts in Translation Data?
Sophie Arnoult and Khalil Sima'an 157