Fourth Workshop on Syntax and Structure in Statistical Translation

(SSST-4)

At the 23rd International Conference on Computational Linguistics 2010

Beijing, China
28 August 2010

Editors:

Dekai Wu

ISBN: 978-1-61782-346-6

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© (2010) by the Chinese Information Processing Society of China
All rights reserved.

Printed by Curran Associates, Inc. (2011)

For permission requests, please contact the Chinese Information Processing Society of China
at the address below.

Chinese Information Processing Society of China
No. 4 Zhongguancun South 4th St.
Haidian District Beijing
China 100190

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

Intersecting Hierarchical and Phrase-Based Models of Translation: Formal Aspects and Algorithms
Marc Dymetman and Nicola Cancedda 1

A Systematic Comparison between Inversion Transduction Grammar and Linear Transduction Grammar for Word Alignment
Markus Saers, Joakim Nivre and Dekai Wu 10

Source-side Syntactic Reordering Patterns with Functional Words for Improved Phrase-based SMT
Jie Jiang, Jinhua Du and Andy Way 19

Syntactic Constraints on Phrase Extraction for Phrase-Based Machine Translation
Hailong Cao, Andrew Finch and Eiichiro Sumita 28

Phrase Based Decoding using a Discriminative Model
Prasanth Kolachina, Sriram Venkatapathy, Srinivas Bangalore, Sudheer Kolachina and Avinesh PVS 34

Seeding Statistical Machine Translation with Translation Memory Output through Tree-Based Structural Alignment
Ventsislav Zhechev and Josef van Genabith 43

Semantic vs. Syntactic vs. N-gram Structure for Machine Translation Evaluation
Chi-kiu Lo and Dekai Wu 52

Arabic morpho-syntactic feature disambiguation in a translation context
Ines Turki Khemakhem, Salma Jamoussi and Abdelmajid Ben hamadou 61

A Discriminative Approach for Dependency Based Statistical Machine Translation
Sriram Venkatapathy, Rajeev Sangal, Aravind Joshi and Karthik Gali 66

Improved Language Modeling for English-Persian Statistical Machine Translation
Mahsa Mohaghegh, Abdolhossein Sarrafzadeh and Tom Moir 75

Manipuri-English Bidirectional Statistical Machine Translation Systems using Morphology and Dependency Relations
Thoudam Doren Singh and Sivaji Bandyopadhyay 83

A Discriminative Syntactic Model for Source Permutation via Tree Transduction
Maxim Khalilov and Khalil Sima'an 92

HMM Word-to-Phrase Alignment with Dependency Constraints
Yanjun Ma and Andy Way 101

New Parameterizations and Features for PSCFG-Based Machine Translation
Andreas Zollmann and Stephan Vogel 110

Deep Syntax Language Models and Statistical Machine Translation
Yvette Graham and Josef van Genabith 118

Conference Program

Saturday, August 28, 2010

9:00–10:40 *Open Tutorial: Tree-Structured and Syntactic SMT*
Dekai Wu

10:40–11:05 Coffee break

11:05–12:05 *Invited Keynote*
Martin Kay

12:05–12:25 *Intersecting Hierarchical and Phrase-Based Models of Translation: Formal Aspects and Algorithms*
Marc Dymetman and Nicola Cancedda

12:25–12:55 *A Systematic Comparison between Inversion Transduction Grammar and Linear Transduction Grammar for Word Alignment*
Markus Saers, Joakim Nivre and Dekai Wu

12:55–14:00 Lunch break

14:00–14:20 *Source-side Syntactic Reordering Patterns with Functional Words for Improved Phrase-based SMT*
Jie Jiang, Jinhua Du and Andy Way

14:20–14:40 *Syntactic Constraints on Phrase Extraction for Phrase-Based Machine Translation*
Hailong Cao, Andrew Finch and Eiichiro Sumita

14:40–15:00 *Phrase Based Decoding using a Discriminative Model*
Prasanth Kolachina, Sriram Venkatapathy, Srinivas Bangalore, Sudheer Kolachina and Avinesh PVS

15:00–15:20 *Seeding Statistical Machine Translation with Translation Memory Output through Tree-Based Structural Alignment*
Ventsislav Zhechev and Josef van Genabith

15:20–15:40 *Semantic vs. Syntactic vs. N-gram Structure for Machine Translation Evaluation*
Chi-kiu Lo and Dekai Wu

15:40–16:05 Posters / Coffee break

Arabic morpho-syntactic feature disambiguation in a translation context
Ines Turki Khemakhem, Salma Jamoussi and Abdelmajid Ben hamadou

Saturday, August 28, 2010 (continued)

A Discriminative Approach for Dependency Based Statistical Machine Translation
Sriram Venkatapathy, Rajeev Sangal, Aravind Joshi and Karthik Gali

Improved Language Modeling for English-Persian Statistical Machine Translation
Mahsa Mohaghegh, Abdolhossein Sarrafzadeh and Tom Moir

Manipuri-English Bidirectional Statistical Machine Translation Systems using Morphology and Dependency Relations
Thoudam Doren Singh and Sivaji Bandyopadhyay

16:05–16:25 *A Discriminative Syntactic Model for Source Permutation via Tree Transduction*
Maxim Khalilov and Khalil Sima'an

16:25–16:45 *HMM Word-to-Phrase Alignment with Dependency Constraints*
Yanjun Ma and Andy Way

16:45–17:05 *New Parameterizations and Features for PSCFG-Based Machine Translation*
Andreas Zollmann and Stephan Vogel

17:05–17:25 *Deep Syntax Language Models and Statistical Machine Translation*
Yvette Graham and Josef van Genabith

17:25–17:40 Discussion