



CONTENTS

Articles

Simple and Efficient Model Filtering in Statistical Machine Translation 5
Juan Pino, Aurelien Waite, William Byrne

Appraise: an Open-Source Toolkit for Manual Evaluation of MT Output 25
Christian Federmann

Hierarchical Phrase-Based Translation with Jane 2 37
Matthias Huck, Jan-Thorsten Peter, Markus Freitag, Stephan Peitz, Hermann Ney

pycdec: A Python Interface to cdec 51
Victor Chahuneau, Noah A. Smith, Chris Dyer

Phrasal Rank-Encoding: Exploiting Phrase Redundancy and Translational Relations for Phrase Table Compression 63
Marcin Junczys-Dowmunt

TrTok: A Fast and Trainable Tokenizer for Natural Languages 75
Jiří Maršík, Ondřej Bojar

Parallel Phrase Scoring for Extra-large Corpora 87
Mohammed Mediani, Jan Niehues, Alex Waibel

rgbF: An Open Source Tool for n-gram Based Automatic Evaluation of Machine Translation Output 99
Maja Popović

Better Splitting Algorithms for Parallel Corpus Processing	109
<i>Lane Schwartz</i>	
DELiC4MT: A Tool for Diagnostic MT Evaluation over User-defined Linguistic Phenomena	121
<i>Antonio Toral, Sudip Kumar Naskar, Federico Gaspari, Declan Groves</i>	
Extending Hiero Decoding in Moses with Cube Growing	133
<i>Wenduan Xu, Philipp Koehn</i>	
Instructions for Authors	143

Articles

2	Simple and Efficient Model Filtering in Statistical Machine Translation	Juan Pino, Aurelien Waks, William Byrne
25	Aggriser: an Open-Source Toolkit for Manual Evaluation of MT Output	Christian Federmann
37	Hierarchical Phrase-Based Translation with Java 5	Mattias Huck, Jan-Thorsten Peter, Markus Freitag, Stephan Fuick, Hermann Ney
51	pydict: A Python Interface to edict	Victor Chahuneau, Noah A. Smith, Chris Dyer
63	Phrase Rank-Encoding: Exploiting Phrase Redundancy and Translational Relations for Phrase Table Compression	Martin Juncjy-Dumont
75	TRUK: A Fast and Flexible Tokenizer for Natural Languages	Jiri Matisk, Ondrej Bojar
87	Parallel Phrase Scoring for Extra-Large Corpora	Mohammed Medani, Jan Niehues, Alex Waibel
99	rgpt: An Open Source Tool for n-gram Based Automatic Evaluation of Machine Translation Output	Maja Popovic