

Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)

Copenhagen, Denmark
7 – 11 September 2017

Volume 1 of 4

ISBN: 978-1-5108-4745-3

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© (2017) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2017)

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: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Monolingual Phrase Alignment on Parse Forests</i> Yuki Arase and Jun'ichi Tsujii	1
<i>Fast(er) Exact Decoding and Global Training for Transition-Based Dependency Parsing via a Minimal Feature Set</i> Tianze Shi, Liang Huang and Lillian Lee	12
<i>Quasi-Second-Order Parsing for 1-Endpoint-Crossing, Pagenumber-2 Graphs</i> Junjie Cao, Sheng Huang, Weiwei Sun and Xiaojun Wan	24
<i>Position-aware Attention and Supervised Data Improve Slot Filling</i> Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli and Christopher D. Manning	35
<i>Heterogeneous Supervision for Relation Extraction: A Representation Learning Approach</i> Liyuan Liu, Xiang Ren, Qi Zhu, Shi Zhi, Huan Gui, Heng Ji and Jiawei Han	46
<i>Integrating Order Information and Event Relation for Script Event Prediction</i> Zhongqing Wang, Yue Zhang and Ching-Yun Chang	57
<i>Entity Linking for Queries by Searching Wikipedia Sentences</i> Chuanqi Tan, Furu Wei, Pengjie Ren, Weifeng Lv and Ming Zhou	68
<i>Train-O-Matic: Large-Scale Supervised Word Sense Disambiguation in Multiple Languages without Manual Training Data</i> Tommaso Pasini and Roberto Navigli	78
<i>Universal Semantic Parsing</i> Siva Reddy, Oscar Täckström, Slav Petrov, Mark Steedman and Mirella Lapata	89
<i>Mimicking Word Embeddings using Subword RNNs</i> Yuval Pinter, Robert Guthrie and Jacob Eisenstein	102
<i>Past, Present, Future: A Computational Investigation of the Typology of Tense in 1000 Languages</i> Ehsaneddin Asgari and Hinrich Schütze	113
<i>Neural Machine Translation with Source-Side Latent Graph Parsing</i> Kazuma Hashimoto and Yoshimasa Tsuruoka	125
<i>Neural Machine Translation with Word Predictions</i> Rongxiang Weng, Shujian Huang, Zaixiang Zheng, XIN-YU DAI and Jiajun CHEN	136
<i>Towards Decoding as Continuous Optimisation in Neural Machine Translation</i> Cong Duy Vu Hoang, Gholamreza Haffari and Trevor Cohn	146
<i>Where is Misty? Interpreting Spatial Descriptors by Modeling Regions in Space</i> Nikita Kitaev and Dan Klein	157
<i>Continuous Representation of Location for Geolocation and Lexical Dialectology using Mixture Density Networks</i> Afshin Rahimi, Timothy Baldwin and Trevor Cohn	167
<i>Obj2Text: Generating Visually Descriptive Language from Object Layouts</i> Xuwang Yin and Vicente Ordonez	177

<i>End-to-end Neural Coreference Resolution</i>	
Kenton Lee, Luheng He, Mike Lewis and Luke Zettlemoyer	188
<i>Neural Net Models of Open-domain Discourse Coherence</i>	
Jiwei Li and Dan Jurafsky	198
<i>Affinity-Preserving Random Walk for Multi-Document Summarization</i>	
Kexiang Wang, Tianyu Liu, Zhifang Sui and Baobao Chang	210
<i>A Mention-Ranking Model for Abstract Anaphora Resolution</i>	
Ana Marasovic, Leo Born, Juri Opitz and Anette Frank	221
<i>Hierarchical Embeddings for Hypernymy Detection and Directionality</i>	
Kim Anh Nguyen, Maximilian Köper, Sabine Schulte im Walde and Ngoc Thang Vu	233
<i>Ngram2vec: Learning Improved Word Representations from Ngram Co-occurrence Statistics</i>	
Zhe Zhao, Tao Liu, Shen Li, Bofang Li and Xiaoyong Du	244
<i>Dict2vec : Learning Word Embeddings using Lexical Dictionaries</i>	
Julien Tissier, Christopher Gravier and Amaury Habrard	254
<i>Learning Chinese Word Representations From Glyphs Of Characters</i>	
Tzu-ray Su and Hung-yi Lee	264
<i>Learning Paraphrastic Sentence Embeddings from Back-Translated Bitext</i>	
John Wieting, Jonathan Mallinson and Kevin Gimpel	274
<i>Joint Embeddings of Chinese Words, Characters, and Fine-grained Subcharacter Components</i>	
Jinxing Yu, Xun Jian, Hao Xin and Yangqiu Song	286
<i>Exploiting Morphological Regularities in Distributional Word Representations</i>	
Arihant Gupta, Syed Sarfaraz Akhtar, Avijit Vajpayee, Arjit Srivastava, Madan Gopal Jhanwar and Manish Shrivastava	292
<i>Exploiting Word Internal Structures for Generic Chinese Sentence Representation</i>	
Shaonan Wang, Jiajun Zhang and Chengqing Zong	298
<i>High-risk learning: acquiring new word vectors from tiny data</i>	
Aurélie Herbelot and Marco Baroni	304
<i>Word Embeddings based on Fixed-Size Ordinally Forgetting Encoding</i>	
Joseph Sanu, Mingbin Xu, Hui Jiang and Quan Liu	310
<i>VecShare: A Framework for Sharing Word Representation Vectors</i>	
Jared Fernandez, Zhaocheng Yu and Doug Downey	316
<i>Word Re-Embedding via Manifold Dimensionality Retention</i>	
Souleiman Hasan and Edward Curry	321
<i>MUSE: Modularizing Unsupervised Sense Embeddings</i>	
Guang-He Lee and Yun-Nung Chen	327
<i>Reporting Score Distributions Makes a Difference: Performance Study of LSTM-networks for Sequence Tagging</i>	
Nils Reimers and Iryna Gurevych	338

<i>Learning What's Easy: Fully Differentiable Neural Easy-First Taggers</i> André F. T. Martins and Julia Kreutzer	349
<i>Incremental Skip-gram Model with Negative Sampling</i> Nobuhiro Kaji and Hayato Kobayashi	363
<i>Learning to select data for transfer learning with Bayesian Optimization</i> Sebastian Ruder and Barbara Plank	372
<i>Unsupervised Pretraining for Sequence to Sequence Learning</i> Prajit Ramachandran, Peter Liu and Quoc Le	383
<i>Efficient Attention using a Fixed-Size Memory Representation</i> Denny Britz, Melody Guan and Minh-Thang Luong	392
<i>Rotated Word Vector Representations and their Interpretability</i> Sungjoon Park, JinYeong Bak and Alice Oh	401
<i>A causal framework for explaining the predictions of black-box sequence-to-sequence models</i> David Alvarez-Melis and Tommi Jaakkola	412
<i>Piecewise Latent Variables for Neural Variational Text Processing</i> Iulian Vlad Serban, Alexander G. Ororbia, Joelle Pineau and Aaron Courville	422
<i>Learning the Structure of Variable-Order CRFs: a finite-state perspective</i> Thomas Lavergne and François Yvon	433
<i>Sparse Communication for Distributed Gradient Descent</i> Alham Fikri Aji and Kenneth Heafield	440
<i>A Joint Many-Task Model: Growing a Neural Network for Multiple NLP Tasks</i> Kazuma Hashimoto, caiming xiong, Yoshimasa Tsuruoka and Richard Socher	446
<i>Why ADAGRAD Fails for Online Topic Modeling</i> You Lu, Jeffrey Lund and Jordan Boyd-Graber	457
<i>Recurrent Attention Network on Memory for Aspect Sentiment Analysis</i> Peng Chen, Zhongqian Sun, Lidong Bing and Wei Yang	463
<i>A Cognition Based Attention Model for Sentiment Analysis</i> Yunfei Long, Lu Qin, Rong Xiang, Minglei Li and Chu-Ren Huang	473
<i>Author-aware Aspect Topic Sentiment Model to Retrieve Supporting Opinions from Reviews</i> Lahari Poddar, Wynne Hsu and Mong Li Lee	483
<i>Magnets for Sarcasm: Making Sarcasm Detection Timely, Contextual and Very Personal</i> Aniruddha Ghosh and Tony Veale	493
<i>Identifying Humor in Reviews using Background Text Sources</i> Alex Morales and Chengxiang Zhai	503
<i>Sentiment Lexicon Construction with Representation Learning Based on Hierarchical Sentiment Supervision</i> Leyi Wang and Rui Xia	513

<i>Towards a Universal Sentiment Classifier in Multiple languages</i> Kui Xu and Xiaojun Wan	522
<i>Capturing User and Product Information for Document Level Sentiment Analysis with Deep Memory Network</i> Zi-Yi Dou	532
<i>Identifying and Tracking Sentiments and Topics from Social Media Texts during Natural Disasters</i> Min Yang, Jincheng Mei, Heng Ji, Zhao Wei, Zhou Zhao and Xiaojun Chen	538
<i>Refining Word Embeddings for Sentiment Analysis</i> Liang-Chih Yu, Jin Wang, K. Robert Lai and Xuejie Zhang	545
<i>A Multilayer Perceptron based Ensemble Technique for Fine-grained Financial Sentiment Analysis</i> Md Shad Akhtar, Abhishek Kumar, Deepanway Ghosal, Asif Ekbal and Pushpak Bhattacharyya	551
<i>Sentiment Intensity Ranking among Adjectives Using Sentiment Bearing Word Embeddings</i> Raksha Sharma, Arpan Somani, Lakshya Kumar and Pushpak Bhattacharyya	558
<i>Sentiment Lexicon Expansion Based on Neural PU Learning, Double Dictionary Lookup, and Polarity Association</i> Yasheng Wang, Yang Zhang and Bing Liu	564
<i>DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning</i> Wenhan Xiong, Thien Hoang and William Yang Wang	575
<i>Task-Oriented Query Reformulation with Reinforcement Learning</i> Rodrigo Nogueira and Kyunghyun Cho	585
<i>Sentence Simplification with Deep Reinforcement Learning</i> Xingxing Zhang and Mirella Lapata	595
<i>Learning how to Active Learn: A Deep Reinforcement Learning Approach</i> Meng Fang, Yuan Li and Trevor Cohn	606
<i>Split and Rephrase</i> Shashi Narayan, Claire Gardent, Shay B. Cohen and Anastasia Shimorina	617
<i>Neural Response Generation via GAN with an Approximate Embedding Layer</i> Zhen Xu, Bingquan Liu, Baoxun Wang, Chengjie SUN, Xiaolong Wang, Zhuoran Wang and Chao Qi	628
<i>A Hybrid Convolutional Variational Autoencoder for Text Generation</i> Stanislau Semeniuta, Aliaksei Severyn and Erhardt Barth	638
<i>Filling the Blanks (hint: plural noun) for Mad Libs Humor</i> Nabil Hossain, John Krumm, Lucy Vanderwende, Eric Horvitz and Henry Kautz	649
<i>Measuring Thematic Fit with Distributional Feature Overlap</i> Enrico Santus, Emmanuele Chersoni, Alessandro Lenci and Philippe Blache	659
<i>SCDV: Sparse Composite Document Vectors using soft clustering over distributional representations</i> Dheeraj Mekala, Vivek Gupta, Bhargavi Paranjape and Harish Karnick	670
<i>Supervised Learning of Universal Sentence Representations from Natural Language Inference Data</i> Alexis Conneau, Douwe Kiela, Holger Schwenk, Loïc Barrault and Antoine Bordes	681

<i>Determining Semantic Textual Similarity using Natural Deduction Proofs</i> Hitomi Yanaka, Koji Mineshima, Pascual Martínez-Gómez and Daisuke Bekki	692
<i>Multi-Grained Chinese Word Segmentation</i> Chen Gong, Zhenghua Li, Min Zhang and Xinzhou Jiang	703
<i>Don't Throw Those Morphological Analyzers Away Just Yet: Neural Morphological Disambiguation for Arabic</i> Nasser Zalmout and Nizar Habash	715
<i>Paradigm Completion for Derivational Morphology</i> Ryan Cotterell, Ekaterina Vylomova, Huda Khayrallah, Christo Kirov and David Yarowsky	725
<i>A Sub-Character Architecture for Korean Language Processing</i> Karl Stratos	732
<i>Do LSTMs really work so well for PoS tagging? – A replication study</i> Tobias Horsmann and Torsten Zesch	738
<i>The Labeled Segmentation of Printed Books</i> Lara McConnaughey, Jennifer Dai and David Bamman	748
<i>Cross-lingual Character-Level Neural Morphological Tagging</i> Ryan Cotterell and Georg Heigold	759
<i>Word-Context Character Embeddings for Chinese Word Segmentation</i> Hao Zhou, Zhenting Yu, Yue Zhang, Shujian Huang, XIN-YU DAI and Jiajun Chen	771
<i>Segmentation-Free Word Embedding for Unsegmented Languages</i> Takamasa Oshikiri	778
<i>From Textbooks to Knowledge: A Case Study in Harvesting Axiomatic Knowledge from Textbooks to Solve Geometry Problems</i> Mrinmaya Sachan, Kumar Dubey and Eric Xing	784
<i>RACE: Large-scale ReAding Comprehension Dataset From Examinations</i> Guokun Lai, Qizhe Xie, Hanxiao Liu, Yiming Yang and Eduard Hovy	796
<i>Beyond Sentential Semantic Parsing: Tackling the Math SAT with a Cascade of Tree Transducers</i> Mark Hopkins, Cristian Petrescu-Prahova, Roie Levin, Ronan Le Bras, Alvaro Herrasti and Vidur Joshi	806
<i>Learning Fine-Grained Expressions to Solve Math Word Problems</i> Danqing Huang, Shuming Shi, Chin-Yew Lin and Jian Yin	816
<i>Structural Embedding of Syntactic Trees for Machine Comprehension</i> Rui Liu, Junjie Hu, Wei Wei, Zi Yang and Eric Nyberg	826
<i>World Knowledge for Reading Comprehension: Rare Entity Prediction with Hierarchical LSTMs Using External Descriptions</i> Teng Long, Emmanuel Bengio, Ryan Lowe, Jackie Chi Kit Cheung and Doina Precup	836
<i>Two-Stage Synthesis Networks for Transfer Learning in Machine Comprehension</i> David Golub, Po-Sen Huang, Xiaodong He and Li Deng	846

<i>Deep Neural Solver for Math Word Problems</i>	
Yan Wang, Xiaojiang Liu and Shuming Shi	856
<i>Latent Space Embedding for Retrieval in Question-Answer Archives</i>	
Deepak P, Dinesh Garg and Shirish Shevade	866
<i>Question Generation for Question Answering</i>	
Nan Duan, Duyu Tang, Peng Chen and Ming Zhou	877
<i>Learning to Paraphrase for Question Answering</i>	
Li Dong, Jonathan Mallinson, Siva Reddy and Mirella Lapata	886
<i>Temporal Information Extraction for Question Answering Using Syntactic Dependencies in an LSTM-based Architecture</i>	
Yuanliang Meng, Anna Rumshisky and Alexey Romanov	898
<i>Ranking Kernels for Structures and Embeddings: A Hybrid Preference and Classification Model</i>	
Kateryna Tymoshenko, Daniele Bonadiman and Alessandro Moschitti	908
<i>Recovering Question Answering Errors via Query Revision</i>	
Semih Yavuz, Izzeddin Gur, Yu Su and Xifeng Yan	914
<i>An empirical study on the effectiveness of images in Multimodal Neural Machine Translation</i>	
Jean-Benoit Delbrouck and Stéphane Dupont	921
<i>Sound-Word2Vec: Learning Word Representations Grounded in Sounds</i>	
Ashwin Vijayakumar, Ramakrishna Vedantam and Devi Parikh	931
<i>The Promise of Premise: Harnessing Question Premises in Visual Question Answering</i>	
Aroma Mahendru, Viraj Prabhu, Akrit Mohapatra, Dhruv Batra and Stefan Lee	937
<i>Guided Open Vocabulary Image Captioning with Constrained Beam Search</i>	
Peter Anderson, Basura Fernando, Mark Johnson and Stephen Gould	947
<i>Zero-Shot Activity Recognition with Verb Attribute Induction</i>	
Rowan Zellers and Yejin Choi	957
<i>Deriving continuous grounded meaning representations from referentially structured multimodal contexts</i>	
Sina Zarrieß and David Schlangen	970
<i>Hierarchically-Attentive RNN for Album Summarization and Storytelling</i>	
Licheng Yu, Mohit Bansal and Tamara Berg	977
<i>Video Highlight Prediction Using Audience Chat Reactions</i>	
Cheng-Yang Fu, Joon Lee, Mohit Bansal and Alexander Berg	983
<i>Reinforced Video Captioning with Entailment Rewards</i>	
Ramakanth Pasunuru and Mohit Bansal	990
<i>Evaluating Hierarchies of Verb Argument Structure with Hierarchical Clustering</i>	
Jesse Mu, Joshua K. Hartshorne and Timothy O'Donnell	997
<i>Incorporating Global Visual Features into Attention-based Neural Machine Translation.</i>	
Iacer Calixto and Qun Liu	1003

<i>Mapping Instructions and Visual Observations to Actions with Reinforcement Learning</i> Dipendra Misra, John Langford and Yoav Artzi	1015
<i>An analysis of eye-movements during reading for the detection of mild cognitive impairment</i> Kathleen C. Fraser, Kristina Lundholm Fors, Dimitrios Kokkinakis and Arto Nordlund	1027
<i>A Structured Learning Approach to Temporal Relation Extraction</i> Qiang Ning, Zhili Feng and Dan Roth	1038
<i>Importance sampling for unbiased on-demand evaluation of knowledge base population</i> Arun Chaganty, Ashwin Paranjape, Percy Liang and Christopher D. Manning	1049
<i>PACRR: A Position-Aware Neural IR Model for Relevance Matching</i> Kai Hui, Andrew Yates, Klaus Berberich and Gerard de Melo	1060
<i>Globally Normalized Reader</i> Jonathan Raiman and John Miller	1070
<i>Speech segmentation with a neural encoder model of working memory</i> Micha Elsner and Cory Shain	1081
<i>Speaking, Seeing, Understanding: Correlating semantic models with conceptual representation in the brain</i> Luana Bulat, Stephen Clark and Ekaterina Shutova	1092
<i>Multi-modal Summarization for Asynchronous Collection of Text, Image, Audio and Video</i> Haoran Li, Junnan Zhu, Cong Ma, Jiajun Zhang and Chengqing Zong	1103
<i>Tensor Fusion Network for Multimodal Sentiment Analysis</i> Amir Zadeh, Minghai Chen, Soujanya Poria, Erik Cambria and Louis-Philippe Morency	1114
<i>ConStance: Modeling Annotation Contexts to Improve Stance Classification</i> Kenneth Joseph, Lisa Friedland, William Hobbs, David Lazer and Oren Tsur	1126
<i>Deeper Attention to Abusive User Content Moderation</i> John Pavlopoulos, Prodromos Malakasiotis and Ion Androutsopoulos	1136
<i>Outta Control: Laws of Semantic Change and Inherent Biases in Word Representation Models</i> Haim Dubossarsky, Daphna Weinshall and Eitan Grossman	1147
<i>Human Centered NLP with User-Factor Adaptation</i> Veronica Lynn, Youngseo Son, Vivek Kulkarni, Niranjan Balasubramanian and H. Andrew Schwartz 1157	
<i>Neural Sequence Learning Models for Word Sense Disambiguation</i> Alessandro Raganato, Claudio Delli Bovi and Roberto Navigli	1167
<i>Learning Word Relatedness over Time</i> Guy D. Rosin, Eytan Adar and Kira Radinsky	1179
<i>Inter-Weighted Alignment Network for Sentence Pair Modeling</i> Gehui Shen, Yunlun Yang and Zhi-Hong Deng	1190
<i>A Short Survey on Taxonomy Learning from Text Corpora: Issues, Resources and Recent Advances</i> Chengyu Wang, Xiaofeng He and Aoying Zhou	1201

<i>Idiom-Aware Compositional Distributed Semantics</i>	
Pengfei Liu, Kaiyu Qian, Xipeng Qiu and Xuanjing Huang	1215
<i>Macro Grammars and Holistic Triggering for Efficient Semantic Parsing</i>	
Yuchen Zhang, Panupong Pasupat and Percy Liang	1225
<i>A Continuously Growing Dataset of Sentential Paraphrases</i>	
Wuwei Lan, Siyu Qiu, Hua He and Wei Xu	1235
<i>Cross-domain Semantic Parsing via Paraphrasing</i>	
Yu Su and Xifeng Yan	1246
<i>A Joint Sequential and Relational Model for Frame-Semantic Parsing</i>	
Bishan Yang and Tom Mitchell	1258
<i>Getting the Most out of AMR Parsing</i>	
Chuan Wang and Nianwen Xue	1268
<i>AMR Parsing using Stack-LSTMs</i>	
Miguel Ballesteros and Yaser Al-Onaizan	1280
<i>An End-to-End Deep Framework for Answer Triggering with a Novel Group-Level Objective</i>	
Jie Zhao, Yu Su, Ziyu Guan and Huan Sun	1287
<i>Predicting Word Association Strengths</i>	
Andrew Cattle and Xiaojuan Ma	1294
<i>Learning Contextually Informed Representations for Linear-Time Discourse Parsing</i>	
Yang Liu and Mirella Lapata	1300
<i>Multi-task Attention-based Neural Networks for Implicit Discourse Relationship Representation and Identification</i>	
Man Lan, Jianxiang Wang, Yuanbin Wu, Zheng-Yu Niu and Haifeng Wang	1310
<i>Chinese Zero Pronoun Resolution with Deep Memory Network</i>	
Qingyu Yin, Yu Zhang, Weinan Zhang and Ting Liu	1320
<i>How much progress have we made on RST discourse parsing? A replication study of recent results on the RST-DT</i>	
Mathieu Morey, Philippe Muller and Nicholas Asher	1330
<i>What is it? Disambiguating the different readings of the pronoun ‘it’</i>	
Sharid Loáiciga, Liane Guillou and Christian Hardmeier	1336
<i>Revisiting Selectional Preferences for Coreference Resolution</i>	
Benjamin Heinzerling, Nafise Sadat Moosavi and Michael Strube	1343
<i>Learning to Rank Semantic Coherence for Topic Segmentation</i>	
Liang Wang, Sujian Li, Yajuan Lv and Houfeng WANG	1351
<i>GRASP: Rich Patterns for Argumentation Mining</i>	
Eyal Shnarch, Ran Levy, Vikas Raykar and Noam Slonim	1356
<i>Patterns of Argumentation Strategies across Topics</i>	
Khalid Al Khatib, Henning Wachsmuth, Matthias Hagen and Benno Stein	1362

<i>Using Argument-based Features to Predict and Analyse Review Helpfulness</i>	
Haijing Liu, Yang Gao, Pin Lv, Mengxue Li, Shiqiang Geng, Minglan Li and Hao Wang	1369
<i>Here's My Point: Joint Pointer Architecture for Argument Mining</i>	
Peter Potash, Alexey Romanov and Anna Rumshisky	1375
<i>Identifying attack and support argumentative relations using deep learning</i>	
Oana Cocarascu and Francesca Toni	1385
<i>Neural Lattice-to-Sequence Models for Uncertain Inputs</i>	
Matthias Sperber, Graham Neubig, Jan Niehues and Alex Waibel	1391
<i>Memory-augmented Neural Machine Translation</i>	
Yang Feng, Shiyue Zhang, Andi Zhang, Dong Wang and Andrew Abel	1401
<i>Dynamic Data Selection for Neural Machine Translation</i>	
Marlies van der Wees, Arianna Bisazza and Christof Monz	1411
<i>Neural Machine Translation Leveraging Phrase-based Models in a Hybrid Search</i>	
Leonard Dahlmann, Evgeny Matusov, Pavel Petrushkov and Shahram Khadivi	1422
<i>Translating Phrases in Neural Machine Translation</i>	
Xing Wang, Zhaopeng Tu, Deyi Xiong and Min Zhang	1432
<i>Towards Bidirectional Hierarchical Representations for Attention-based Neural Machine Translation</i>	
Baosong Yang, Derek F. Wong, Tong Xiao, Lidia S. Chao and Jingbo Zhu	1443
<i>Learning Translations via Matrix Completion</i>	
Derry Tanti Wijaya, Brendan Callahan, John Hewitt, Jie Gao, Xiao Ling, Marianna Apidianaki and Chris Callison-Burch	1453
<i>Reinforcement Learning for Bandit Neural Machine Translation with Simulated Human Feedback</i>	
Khanh Nguyen, Hal Daumé III and Jordan Boyd-Graber	1465
<i>Towards Compact and Fast Neural Machine Translation Using a Combined Method</i>	
Xiaowei Zhang, Wei Chen, Feng Wang, Shuang Xu and Bo Xu	1476
<i>Instance Weighting for Neural Machine Translation Domain Adaptation</i>	
Rui Wang, Masao Utiyama, Lemao Liu, Kehai Chen and Eiichiro Sumita	1483
<i>Regularization techniques for fine-tuning in neural machine translation</i>	
Antonio Valerio Miceli Barone, Barry Haddow, Ulrich Germann and Rico Sennrich	1490
<i>Source-Side Left-to-Right or Target-Side Left-to-Right? An Empirical Comparison of Two Phrase-Based Decoding Algorithms</i>	
Yin-Wen Chang and Michael Collins	1496
<i>Using Target-side Monolingual Data for Neural Machine Translation through Multi-task Learning</i>	
Tobias Domhan and Felix Hieber	1501
<i>Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling</i>	
Diego Marcheggiani and Ivan Titov	1507
<i>Neural Semantic Parsing with Type Constraints for Semi-Structured Tables</i>	
Jayant Krishnamurthy, Pradeep Dasigi and Matt Gardner	1517

<i>Joint Concept Learning and Semantic Parsing from Natural Language Explanations</i> Shashank Srivastava, Igor Labutov and Tom Mitchell	1528
<i>Grasping the Finer Point: A Supervised Similarity Network for Metaphor Detection</i> Marek Rei, Luana Bulat, Douwe Kiela and Ekaterina Shutova	1538
<i>Identifying civilians killed by police with distantly supervised entity-event extraction</i> Katherine Keith, Abram Handler, Michael Pinkham, Cara Magliozzi, Joshua McDuffie and Brendan O'Connor	1548
<i>Asking too much? The rhetorical role of questions in political discourse</i> Justine Zhang, Arthur Spirling and Cristian Danescu-Niculescu-Mizil	1559
<i>Detecting Perspectives in Political Debates</i> David Vilares and Yulan He	1574
<i>"i have a feeling trump will win.....": Forecasting Winners and Losers from User Predictions on Twitter</i> Sandesh Swamy, Alan Ritter and Marie-Catherine de Marneffe	1584
<i>A Question Answering Approach for Emotion Cause Extraction</i> Lin Gui, Jiannan Hu, Yulan He, Ruifeng Xu, Lu Qin and Jiachen Du	1594
<i>Story Comprehension for Predicting What Happens Next</i> Snigdha Chaturvedi, Haoruo Peng and Dan Roth	1604
<i>Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm</i> Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan and Sune Lehmann	1616
<i>Opinion Recommendation Using A Neural Model</i> Zhongqing Wang and Yue Zhang	1627
<i>CRF Autoencoder for Unsupervised Dependency Parsing</i> Jiong Cai, Yong Jiang and Kewei Tu	1639
<i>Efficient Discontinuous Phrase-Structure Parsing via the Generalized Maximum Spanning Arborescence</i> Caio Corro, Joseph Le Roux and Mathieu Lacroix	1645
<i>Incremental Graph-based Neural Dependency Parsing</i> Xiaoqing Zheng	1656
<i>Neural Discontinuous Constituency Parsing</i> Miloš Stanojević and Raquel Garrido Alhama	1667
<i>Stack-based Multi-layer Attention for Transition-based Dependency Parsing</i> Zhirui Zhang, Shujie Liu, Mu Li, Ming Zhou and Enhong Chen	1678
<i>Dependency Grammar Induction with Neural Lexicalization and Big Training Data</i> Wenjuan Han, Yong Jiang and Kewei Tu	1684
<i>Combining Generative and Discriminative Approaches to Unsupervised Dependency Parsing via Dual Decomposition</i> Yong Jiang, Wenjuan Han and Kewei Tu	1690

<i>Effective Inference for Generative Neural Parsing</i>	
Mitchell Stern, Daniel Fried and Dan Klein	1696
<i>Semi-supervised Structured Prediction with Neural CRF Autoencoder</i>	
Xiao Zhang, Yong Jiang, Hao Peng, Kewei Tu and Dan Goldwasser	1702
<i>TAG Parsing with Neural Networks and Vector Representations of Supertags</i>	
Jungo Kasai, Bob Frank, Tom McCoy, Owen Rambow and Alexis Nasr	1713
<i>Global Normalization of Convolutional Neural Networks for Joint Entity and Relation Classification</i>	
Heike Adel and Hinrich Schütze	1724
<i>End-to-End Neural Relation Extraction with Global Optimization</i>	
Meishan Zhang, Yue Zhang and Guohong Fu	1731
<i>KGEval: Accuracy Estimation of Automatically Constructed Knowledge Graphs</i>	
Prakhar Ojha and Partha Talukdar	1742
<i>Sparsity and Noise: Where Knowledge Graph Embeddings Fall Short</i>	
Jay Pujara, Eriq Augustine and Lise Getoor	1752
<i>Dual Tensor Model for Detecting Asymmetric Lexico-Semantic Relations</i>	
Goran Glavaš and Simone Paolo Ponzetto	1758
<i>Incorporating Relation Paths in Neural Relation Extraction</i>	
Wenyuan Zeng, Yankai Lin, Zhiyuan Liu and Maosong Sun	1769
<i>Adversarial Training for Relation Extraction</i>	
Yi Wu, David Bamman and Stuart Russell	1779
<i>Context-Aware Representations for Knowledge Base Relation Extraction</i>	
Daniil Sorokin and Iryna Gurevych	1785
<i>A Soft-label Method for Noise-tolerant Distantly Supervised Relation Extraction</i>	
Tianyu Liu, Kexiang Wang, Baobao Chang and Zhifang Sui	1791
<i>A Sequential Model for Classifying Temporal Relations between Intra-Sentence Events</i>	
Prafulla Kumar Choubey and Ruihong Huang	1797
<i>Deep Residual Learning for Weakly-Supervised Relation Extraction</i>	
Yi Yao Huang and William Yang Wang	1804
<i>Noise-Clustered Distant Supervision for Relation Extraction: A Nonparametric Bayesian Perspective</i>	
Qing Zhang and Houfeng Wang	1809
<i>Exploring Vector Spaces for Semantic Relations</i>	
Kata Gábor, Haifa Zargayouna, Isabelle Tellier, Davide Buscaldi and Thierry Charnois	1815
<i>Temporal dynamics of semantic relations in word embeddings: an application to predicting armed conflict participants</i>	
Andrey Kutuzov, Erik Velldal and Lilja Øvreid	1825
<i>Dynamic Entity Representations in Neural Language Models</i>	
Yangfeng Ji, Chenhao Tan, Sebastian Martschat, Yejin Choi and Noah A. Smith	1831

<i>Towards Quantum Language Models</i>	
Ivano Basile and Fabio Tamburini	1841
<i>Reference-Aware Language Models</i>	
Zichao Yang, Phil Blunsom, Chris Dyer and Wang Ling	1851
<i>A Simple Language Model based on PMI Matrix Approximations</i>	
Oren Melamud, Ido Dagan and Jacob Goldberger	1861
<i>Syllable-aware Neural Language Models: A Failure to Beat Character-aware Ones</i>	
Zhenisbek Assylbekov, Rustem Takhanov, Bagdat Myrzakhmetov and Jonathan N. Washington	
1867	
<i>Inducing Semantic Micro-Clusters from Deep Multi-View Representations of Novels</i>	
Lea Frermann and György Szarvas	1874
<i>Initializing Convolutional Filters with Semantic Features for Text Classification</i>	
Shen Li, Zhe Zhao, Tao Liu, Renfen Hu and Xiaoyong Du	1885
<i>Shortest-Path Graph Kernels for Document Similarity</i>	
Giannis Nikolentzos, Polykarpos Meladianos, Francois Rousseau, Yannis Stavrakas and Michalis Vazirgiannis	1891
<i>Adapting Topic Models using Lexical Associations with Tree Priors</i>	
Weiwei Yang, Jordan Boyd-Graber and Philip Resnik	1902
<i>Finding Patterns in Noisy Crowds: Regression-based Annotation Aggregation for Crowdsourced Data</i>	
Natalie Parde and Rodney Nielsen	1908
<i>CROWD-IN-THE-LOOP: A Hybrid Approach for Annotating Semantic Roles</i>	
Chenguang Wang, Alan Akbik, Iulia Chiticariu, Yunyao Li, Fei Xia and Anbang Xu	1914
<i>Earth Mover's Distance Minimization for Unsupervised Bilingual Lexicon Induction</i>	
Meng Zhang, Yang Liu, Huanbo Luan and Maosong Sun	1924
<i>Unfolding and Shrinking Neural Machine Translation Ensembles</i>	
Felix Stahlberg and Bill Byrne	1936
<i>Graph Convolutional Encoders for Syntax-aware Neural Machine Translation</i>	
Joost Bastings, Ivan Titov, Wilker Aziz, Diego Marcheggiani and Khalil Simaan	1947
<i>Trainable Greedy Decoding for Neural Machine Translation</i>	
Jiatao Gu, Kyunghyun Cho and Victor O.K. Li	1958
<i>Satirical News Detection and Analysis using Attention Mechanism and Linguistic Features</i>	
Fan Yang, Arjun Mukherjee and Eduard Dragut	1969
<i>Fine Grained Citation Span for References in Wikipedia</i>	
Besnik Fetahu, Katja Markert and Avishek Anand	1980
<i>Identifying Semantic Edit Intentions from Revisions in Wikipedia</i>	
Diyi Yang, Aaron Halfaker, Robert Kraut and Eduard Hovy	1990
<i>Accurate Supervised and Semi-Supervised Machine Reading for Long Documents</i>	
Daniel Hewlett, Llion Jones, Alexandre Lacoste and izzeddin gur	2001

<i>Adversarial Examples for Evaluating Reading Comprehension Systems</i>	
Robin Jia and Percy Liang	2011
<i>Reasoning with Heterogeneous Knowledge for Commonsense Machine Comprehension</i>	
Hongyu Lin, Le Sun and Xianpei Han	2022
<i>Document-Level Multi-Aspect Sentiment Classification as Machine Comprehension</i>	
Yichun Yin, Yangqiu Song and Ming Zhang	2034
<i>What is the Essence of a Claim? Cross-Domain Claim Identification</i>	
Johannes Daxenberger, Steffen Eger, Ivan Habernal, Christian Stab and Iryna Gurevych	2045
<i>Identifying Where to Focus in Reading Comprehension for Neural Question Generation</i>	
Xinya Du and Claire Cardie	2057
<i>Break it Down for Me: A Study in Automated Lyric Annotation</i>	
Lucas Sterckx, Jason Naradowsky, Bill Byrne, Thomas Demeester and Chris Develder	2064
<i>Cascaded Attention based Unsupervised Information Distillation for Compressive Summarization</i>	
Piji Li, Wai Lam, Lidong Bing, Weiwei Guo and Hang Li	2071
<i>Deep Recurrent Generative Decoder for Abstractive Text Summarization</i>	
Piji Li, Wai Lam, Lidong Bing and Zihao Wang	2081
<i>Extractive Summarization Using Multi-Task Learning with Document Classification</i>	
Masaru Isonuma, Toru Fujino, Junichiro Mori, Yutaka Matsuo and Ichiro Sakata	2091
<i>Towards Automatic Construction of News Overview Articles by News Synthesis</i>	
Jianmin Zhang and Xiaojun Wan	2101
<i>Joint Syntacto-Discourse Parsing and the Syntacto-Discourse Treebank</i>	
Kai Zhao and Liang Huang	2107
<i>Event Coreference Resolution by Iteratively Unfolding Inter-dependencies among Events</i>	
Prafulla Kumar Choubey and Ruihong Huang	2114
<i>When to Finish? Optimal Beam Search for Neural Text Generation (modulo beam size)</i>	
Liang Huang, Kai Zhao and Mingbo Ma	2124
<i>Steering Output Style and Topic in Neural Response Generation</i>	
Di Wang, Nebojsa Jojic, Chris Brockett and Eric Nyberg	2130
<i>Preserving Distributional Information in Dialogue Act Classification</i>	
Quan Hung Tran, Ingrid Zukerman and Gholamreza Haffari	2141
<i>Adversarial Learning for Neural Dialogue Generation</i>	
Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter and Dan Jurafsky	2147
<i>Using Context Information for Dialog Act Classification in DNN Framework</i>	
Yang Liu, Kun Han, Zhao Tan and Yun Lei	2160
<i>Modeling Dialogue Acts with Content Word Filtering and Speaker Preferences</i>	
Yohan Jo, Michael Yoder, Hyeju Jang and Carolyn Rose	2169
<i>Towards Implicit Content-Introducing for Generative Short-Text Conversation Systems</i>	
Lili Yao, Yaoyuan Zhang, Yansong Feng, Dongyan Zhao and Rui Yan	2180

<i>Affordable On-line Dialogue Policy Learning</i>	
Cheng Chang, Runzhe Yang, Lu Chen, Xiang Zhou and Kai Yu	2190
<i>Generating High-Quality and Informative Conversation Responses with Sequence-to-Sequence Models</i>	
Yuanlong Shao, Stephan Gouws, Denny Britz, Anna Goldie, Brian Strope and Ray Kurzweil	2200
<i>Bootstrapping incremental dialogue systems from minimal data: the generalisation power of dialogue grammars</i>	
Arash Eshghi, Igor Shalyminov and Oliver Lemon	2210
<i>Composite Task-Completion Dialogue Policy Learning via Hierarchical Deep Reinforcement Learning</i>	
Baolin Peng, Xiujun Li, Lihong Li, Jianfeng Gao, Asli Celikyilmaz, Sungjin Lee and Kam-Fai Wong	2221
<i>Why We Need New Evaluation Metrics for NLG</i>	
Jekaterina Novikova, Ondřej Dušek, Amanda Cercas Curry and Verena Rieser	2231
<i>Challenges in Data-to-Document Generation</i>	
Sam Wiseman, Stuart Shieber and Alexander Rush	2243
<i>All that is English may be Hindi: Enhancing language identification through automatic ranking of the likeliness of word borrowing in social media</i>	
Jasabanta Patro, Bidisha Samanta, Saurabh Singh, Abhipsa Basu, Prithwish Mukherjee, Monojit Choudhury and Animesh Mukherjee	2254
<i>Multi-View Unsupervised User Feature Embedding for Social Media-based Substance Use Prediction</i>	
Tao Ding, Warren K. Bickel and Shimei Pan	2265
<i>Demographic-aware word associations</i>	
Aparna Garimella, Carmen Banea and Rada Mihalcea	2275
<i>A Factored Neural Network Model for Characterizing Online Discussions in Vector Space</i>	
Hao Cheng, Hao Fang and Mari Ostendorf	2286
<i>Dimensions of Interpersonal Relationships: Corpus and Experiments</i>	
Farzana Rashid and Eduardo Blanco	2297
<i>Argument Mining on Twitter: Arguments, Facts and Sources</i>	
Mihai Dusmanu, Elena Cabrio and Serena Villata	2307
<i>Distinguishing Japanese Non-standard Usages from Standard Ones</i>	
Tatsuya Aoki, Ryohei Sasano, Hiroya Takamura and Manabu Okumura	2313
<i>Connotation Frames of Power and Agency in Modern Films</i>	
Maarten Sap, Marcella Cindy Prasettio, Ari Holtzman, Hannah Rashkin and Yejin Choi	2319
<i>Controlling Human Perception of Basic User Traits</i>	
Daniel Preotjiuc-Pietro, Sharath Chandra Guntuku and Lyle Ungar	2325
<i>Topic Signatures in Political Campaign Speeches</i>	
Clément Gautrais, Peggy Cellier, René Quiniou and Alexandre Termier	2332
<i>Assessing Objective Recommendation Quality through Political Forecasting</i>	
H. Andrew Schwartz, Masoud Rouhizadeh, Michael Bishop, Philip Tetlock, Barbara Mellers and Lyle Ungar	2338

<i>Never Abandon Minorities: Exhaustive Extraction of Bursty Phrases on Microblogs Using Set Cover Problem</i>	
Masumi Shirakawa, Takahiro Hara and Takuya Maekawa	2348
<i>Maximum Margin Reward Networks for Learning from Explicit and Implicit Supervision</i>	
Haoruo Peng, Ming-Wei Chang and Wen-tau Yih	2358
<i>The Impact of Modeling Overall Argumentation with Tree Kernels</i>	
Henning Wachsmuth, Giovanni Da San Martino, Dora Kiesel and Benno Stein	2369
<i>Learning Generic Sentence Representations Using Convolutional Neural Networks</i>	
Zhe Gan, Yunchen Pu, Ricardo Henao, Chunyuan Li, Xiaodong He and Lawrence Carin	2380
<i>Repeat before Forgetting: Spaced Repetition for Efficient and Effective Training of Neural Networks</i>	
Hadi Amiri, Timothy Miller and Guergana Savova	2391
<i>Part-of-Speech Tagging for Twitter with Adversarial Neural Networks</i>	
Tao Gui, Qi Zhang, Haoran Huang, Minlong Peng and Xuanjing Huang	2401
<i>Investigating Different Syntactic Context Types and Context Representations for Learning Word Embeddings</i>	
Bofang Li, Tao Liu, Zhe Zhao, Buzhou Tang, Aleksandr Drozd, Anna Rogers and Xiaoyong Du	2411
<i>Does syntax help discourse segmentation? Not so much</i>	
Chloé Braud, Ophélie Lacroix and Anders Søgaard	2422
<i>Deal or No Deal? End-to-End Learning of Negotiation Dialogues</i>	
Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh and Dhruv Batra	2433
<i>Agent-Aware Dropout DQN for Safe and Efficient On-line Dialogue Policy Learning</i>	
Lu Chen, Xiang Zhou, Cheng Chang, Runzhe Yang and Kai Yu	2444
<i>Towards Debate Automation: a Recurrent Model for Predicting Debate Winners</i>	
Peter Potash and Anna Rumshisky	2455
<i>Further Investigation into Reference Bias in Monolingual Evaluation of Machine Translation</i>	
Qingsong Ma, Yvette Graham, Timothy Baldwin and Qun Liu	2466
<i>A Challenge Set Approach to Evaluating Machine Translation</i>	
Pierre Isabelle, Colin Cherry and George Foster	2476
<i>Knowledge Distillation for Bilingual Dictionary Induction</i>	
Ndapandula Nakashole and Raphael Flauger	2487
<i>Machine Translation, it's a question of style, inmit? The case of English tag questions</i>	
Rachel Bawden	2497
<i>Deciphering Related Languages</i>	
Nima Pourdamghani and Kevin Knight	2503
<i>Identifying Cognate Sets Across Dictionaries of Related Languages</i>	
Adam St Arnaud, David Beck and Grzegorz Kondrak	2509
<i>Learning Language Representations for Typology Prediction</i>	
Chaitanya Malaviya, Graham Neubig and Patrick Littell	2519

<i>Cheap Translation for Cross-Lingual Named Entity Recognition</i> Stephen Mayhew, Chen-Tse Tsai and Dan Roth	2526
<i>Cross-Lingual Induction and Transfer of Verb Classes Based on Word Vector Space Specialisation</i> Ivan Vulić, Nikola Mrkšić and Anna Korhonen	2536
<i>Classification of telicity using cross-linguistic annotation projection</i> Annemarie Friedrich and Damyana Gateva	2549
<i>Counterfactual Learning from Bandit Feedback under Deterministic Logging : A Case Study in Statistical Machine Translation</i> Carolin Lawrence, Artem Sokolov and Stefan Riezler	2556
<i>Learning Fine-grained Relations from Chinese User Generated Categories</i> Chengyu Wang, Yan Fan, Xiaofeng He and Aoying Zhou	2567
<i>Improving Slot Filling Performance with Attentive Neural Networks on Dependency Structures</i> Lifu Huang, Avirup Sil, Heng Ji and Radu Florian	2578
<i>Identifying Products in Online Cybercrime Marketplaces: A Dataset for Fine-grained Domain Adaptation</i> Greg Durrett, Jonathan K. Kummerfeld, Taylor Berg-Kirkpatrick, Rebecca Portnoff, Sadia Afroz, Damon McCoy, Kirill Levchenko and Vern Paxson	2588
<i>Labeling Gaps Between Words: Recognizing Overlapping Mentions with Mention Separators</i> Aldrian Obaja Muis and Wei Lu	2598
<i>Deep Joint Entity Disambiguation with Local Neural Attention</i> Octavian-Eugen Ganea and Thomas Hofmann	2609
<i>MinIE: Minimizing Facts in Open Information Extraction</i> Kiril Gashteovski, Rainer Gemulla and Luciano Del Corro	2620
<i>Scientific Information Extraction with Semi-supervised Neural Tagging</i> Yi Luan, Mari Ostendorf and Hannaneh Hajishirzi	2631
<i>NITE: A Neural Inductive Teaching Framework for Domain Specific NER</i> Siliang Tang, Ning Zhang, Jinjiang Zhang, Fei Wu and Yueting Zhuang	2642
<i>Speeding up Reinforcement Learning-based Information Extraction Training using Asynchronous Methods</i> Aditya Sharma, Zarana Parekh and Partha Talukdar	2648
<i>Leveraging Linguistic Structures for Named Entity Recognition with Bidirectional Recursive Neural Networks</i> Peng-Hsuan Li, Ruo-Ping Dong, Yu-Siang Wang, Ju-Chieh Chou and Wei-Yun Ma	2654
<i>Fast and Accurate Entity Recognition with Iterated Dilated Convolutions</i> Emma Strubell, Patrick Verga, David Belanger and Andrew McCallum	2660
<i>Entity Linking via Joint Encoding of Types, Descriptions, and Context</i> Nitish Gupta, Sameer Singh and Dan Roth	2671
<i>An Insight Extraction System on BioMedical Literature with Deep Neural Networks</i> Hua He, Kris Ganjam, Navendu Jain, Jessica Lundin, Ryen White and Jimmy Lin	2681

<i>Word Etymology as Native Language Interference</i>	
Vivi Nastase and Carlo Strapparava	2692
<i>A Simpler and More Generalizable Story Detector using Verb and Character Features</i>	
Joshua Eisenberg and Mark Finlayson	2698
<i>Multi-modular domain-tailored OCR post-correction</i>	
Sarah Schulz and Jonas Kuhn	2706
<i>Learning to Predict Charges for Criminal Cases with Legal Basis</i>	
Bingfeng Luo, Yansong Feng, Jianbo Xu, Xiang Zhang and Dongyan Zhao	2717
<i>Quantifying the Effects of Text Duplication on Semantic Models</i>	
Alexandra Schofield, Laure Thompson and David Mimno	2727
<i>Identifying Semantically Deviating Outlier Documents</i>	
Honglei Zhuang, Chi Wang, Fangbo Tao, Lance Kaplan and Jiawei Han	2738
<i>Detecting and Explaining Causes From Text For a Time Series Event</i>	
Dongyeop Kang, Varun Gangal, Ang Lu, Zheng Chen and Eduard Hovy	2748
<i>A Novel Cascade Model for Learning Latent Similarity from Heterogeneous Sequential Data of MOOC</i>	
Zhuoxuan Jiang, Shanshan Feng, Gao Cong, Chunyan Miao and Xiaoming Li	2758
<i>Identifying the Provision of Choices in Privacy Policy Text</i>	
Kanthashree Mysore Sathyendra, Shomir Wilson, Florian Schaub, Sebastian Zimmeck and Norman Sadeh	2764
<i>An Empirical Analysis of Edit Importance between Document Versions</i>	
Tanya Goyal, Sachin Kelkar, Manas Agarwal and Jeenu Grover	2770
<i>Transition-Based Disfluency Detection using LSTMs</i>	
Shaolei Wang, Wanxiang Che, Yue Zhang, Meishan Zhang and Ting Liu	2775
<i>Neural Sequence-Labelling Models for Grammatical Error Correction</i>	
Helen Yannakoudakis, Marek Rei, Øistein E. Andersen and Zheng Yuan	2785
<i>Adapting Sequence Models for Sentence Correction</i>	
Allen Schmalz, Yoon Kim, Alexander Rush and Stuart Shieber	2797
<i>A Study of Style in Machine Translation: Controlling the Formality of Machine Translation Output</i>	
Xing Niu, Marianna Martindale and Marine Carpuat	2804
<i>Sharp Models on Dull Hardware: Fast and Accurate Neural Machine Translation Decoding on the CPU</i>	
Jacob Devlin	2810
<i>Exploiting Cross-Sentence Context for Neural Machine Translation</i>	
Longyue Wang, Zhaopeng Tu, Andy Way and Qun Liu	2816
<i>Cross-Lingual Transfer Learning for POS Tagging without Cross-Lingual Resources</i>	
Joo-Kyung Kim, Young-Bum Kim, Ruhi Sarikaya and Eric Fosler-Lussier	2822
<i>Image Pivoting for Learning Multilingual Multimodal Representations</i>	
Spandana Gella, Rico Sennrich, Frank Keller and Mirella Lapata	2829

<i>Neural Machine Translation with Source Dependency Representation</i> Kehai Chen, Rui Wang, Masao Utiyama, Lemao Liu, Akihiro Tamura, Eiichiro Sumita and Tiejun Zhao	2836
<i>Visual Denotations for Recognizing Textual Entailment</i> Dan Han, Pascual Martínez-Gómez and Koji Mineshima	2843
<i>Sequence Effects in Crowdsourced Annotations</i> Nitika Mathur, Timothy Baldwin and Trevor Cohn	2850
<i>No Need to Pay Attention: Simple Recurrent Neural Networks Work!</i> Ferhan Ture and Oliver Jojic	2856
<i>The strange geometry of skip-gram with negative sampling</i> David Mimno and Laure Thompson	2863
<i>Natural Language Processing with Small Feed-Forward Networks</i> Jan A. Botha, Emily Pitler, Ji Ma, Anton Bakalov, Alex Salcianu, David Weiss, Ryan McDonald and Slav Petrov	2869
<i>Deep Multi-Task Learning for Aspect Term Extraction with Memory Interaction</i> Xin Li and Wai Lam	2876
<i>Analogs of Linguistic Structure in Deep Representations</i> Jacob Andreas and Dan Klein	2883
<i>A Simple Regularization-based Algorithm for Learning Cross-Domain Word Embeddings</i> Wei Yang, Wei Lu and Vincent Zheng	2888
<i>Learning what to read: Focused machine reading</i> Enrique Noriega-Atala, Marco A. Valenzuela-Escárcega, Clayton Morrison and Mihai Surdeanu	2895
<i>DOC: Deep Open Classification of Text Documents</i> Lei Shu, Hu Xu and Bing Liu	2901
<i>Charmanteau: Character Embedding Models For Portmanteau Creation</i> Varun Gangal, Harsh Jhamtani, Graham Neubig, Eduard Hovy and Eric Nyberg	2907
<i>Using Automated Metaphor Identification to Aid in Detection and Prediction of First-Episode Schizophrenia</i> E. Dario Gutierrez, Guillermo Cecchi, Cheryl Corcoran and Philip Corlett	2913
<i>Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking</i> Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova and Yejin Choi	2921
<i>Topic-Based Agreement and Disagreement in US Electoral Manifestos</i> Stefano Menini, Federico Nanni, Simone Paolo Ponzetto and Sara Tonelli	2928
<i>Zipporah: a Fast and Scalable Data Cleaning System for Noisy Web-Crawled Parallel Corpora</i> Hainan Xu and Philipp Koehn	2935
<i>Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints</i> Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez and Kai-Wei Chang	2941

<i>Natural Language Does Not Emerge 'Naturally' in Multi-Agent Dialog</i>	
Satwik Kottur, José Moura, Stefan Lee and Dhruv Batra	2952
<i>Depression and Self-Harm Risk Assessment in Online Forums</i>	
Andrew Yates, Arman Cohan and Nazli Goharian	2958
<i>Bringing Structure into Summaries: Crowdsourcing a Benchmark Corpus of Concept Maps</i>	
Tobias Falke and Iryna Gurevych	2969