

# **Conference on Empirical Methods in Natural Language Processing (EMNLP 2018)**

Brussels, Belgium  
31 October - 4 November 2018

Volume 1 of 7

ISBN: 978-1-5108-7416-9

**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© (2018) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2019)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Privacy-preserving Neural Representations of Text</i> Maximin Coavoux, Shashi Narayan and Shay B. Cohen . . . . .	1
<i>Adversarial Removal of Demographic Attributes from Text Data</i> Yanai Elazar and Yoav Goldberg . . . . .	11
<i>DeClarE: Debunking Fake News and False Claims using Evidence-Aware Deep Learning</i> Kashyap Papat, Subhabrata Mukherjee, Andrew Yates and Gerhard Weikum . . . . .	22
<i>It's going to be okay: Measuring Access to Support in Online Communities</i> Zijian Wang and David Jurgens . . . . .	33
<i>Detecting Gang-Involved Escalation on Social Media Using Context</i> Serina Chang, Ruiqi Zhong, Ethan Adams, Fei-Tzin Lee, Siddharth Varia, Desmond Patton, William Frey, Chris Kedzie and Kathy McKeown . . . . .	46
<i>Reasoning about Actions and State Changes by Injecting Commonsense Knowledge</i> Niket Tandon, Bhavana Dalvi, Joel Grus, Wen-tau Yih, Antoine Bosselut and Peter Clark . . . . .	57
<i>Collecting Diverse Natural Language Inference Problems for Sentence Representation Evaluation</i> Adam Poliak, Aparajita Haldar, Rachel Rudinger, J. Edward Hu, Ellie Pavlick, Aaron Steven White and Benjamin Van Durme . . . . .	67
<i>Textual Analogy Parsing: What's Shared and What's Compared among Analogous Facts</i> Matthew Lamm, Arun Chaganty, Christopher D. Manning, Dan Jurafsky and Percy Liang . . . . .	82
<i>SWAG: A Large-Scale Adversarial Dataset for Grounded Commonsense Inference</i> Rowan Zellers, Yonatan Bisk, Roy Schwartz and Yejin Choi . . . . .	93
<i>TwoWingOS: A Two-Wing Optimization Strategy for Evidential Claim Verification</i> Wenpeng Yin and Dan Roth . . . . .	105
<i>Associative Multichannel Autoencoder for Multimodal Word Representation</i> Shaonan Wang, Jiajun Zhang and Chengqing Zong . . . . .	115
<i>Game-Based Video-Context Dialogue</i> Ramakanth Pasunuru and Mohit Bansal . . . . .	125
<i>simNet: Stepwise Image-Topic Merging Network for Generating Detailed and Comprehensive Image Captions</i> Fenglin Liu, Xuancheng Ren, Yuanxin Liu, Houfeng Wang and Xu Sun . . . . .	137
<i>Multimodal Language Analysis with Recurrent Multistage Fusion</i> Paul Pu Liang, Ziyin Liu, AmirAli Bagher Zadeh and Louis-Philippe Morency . . . . .	150
<i>Temporally Grounding Natural Sentence in Video</i> Jingyuan Chen, Xinpeng Chen, Lin Ma, Zequn Jie and Tat-Seng Chua . . . . .	162
<i>PreCo: A Large-scale Dataset in Preschool Vocabulary for Coreference Resolution</i> Hong Chen, Zhenhua Fan, Hao Lu, Alan Yuille and Shu Rong . . . . .	172
<i>Adversarial Transfer Learning for Chinese Named Entity Recognition with Self-Attention Mechanism</i> Pengfei Cao, Yubo Chen, Kang Liu, Jun Zhao and Shengping Liu . . . . .	182

<i>Using Linguistic Features to Improve the Generalization Capability of Neural Coreference Resolvers</i> Nafise Sadat Moosavi and Michael Strube .....	193
<i>Neural Segmental Hypergraphs for Overlapping Mention Recognition</i> Bailin Wang and Wei Lu .....	204
<i>Variational Sequential Labelers for Semi-Supervised Learning</i> Mingda Chen, Qingming Tang, Karen Livescu and Kevin Gimpel .....	215
<i>Joint Representation Learning of Cross-lingual Words and Entities via Attentive Distant Supervision</i> Yixin Cao, Lei Hou, Juanzi Li, Zhiyuan Liu, Chengjiang Li, Xu Chen and Tiansi Dong .....	227
<i>Deep Pivot-Based Modeling for Cross-language Cross-domain Transfer with Minimal Guidance</i> Yftah Ziser and Roi Reichart .....	238
<i>Multi-lingual Common Semantic Space Construction via Cluster-consistent Word Embedding</i> Lifu Huang, Kyunghyun Cho, Boliang Zhang, Heng Ji and Kevin Knight .....	250
<i>Unsupervised Multilingual Word Embeddings</i> Xilun Chen and Claire Cardie .....	261
<i>CLUSE: Cross-Lingual Unsupervised Sense Embeddings</i> Ta Chung Chi and Yun-Nung Chen .....	271
<i>Adversarial Propagation and Zero-Shot Cross-Lingual Transfer of Word Vector Specialization</i> Edoardo Maria Ponti, Ivan Vulić, Goran Glavaš, Nikola Mrkšić and Anna Korhonen .....	282
<i>Improving Cross-Lingual Word Embeddings by Meeting in the Middle</i> Yerai Doval, Jose Camacho-Collados, Luis Espinosa Anke and Steven Schockaert .....	294
<i>WikiAtomicEdits: A Multilingual Corpus of Wikipedia Edits for Modeling Language and Discourse</i> Manaal Faruqui, Ellie Pavlick, Ian Tenney and Dipanjan Das .....	305
<i>On the Relation between Linguistic Typology and (Limitations of) Multilingual Language Modeling</i> Daniela Gerz, Ivan Vulić, Edoardo Maria Ponti, Roi Reichart and Anna Korhonen .....	316
<i>A Fast, Compact, Accurate Model for Language Identification of Codemixed Text</i> Yuan Zhang, Jason Riesa, Daniel Gillick, Anton Bakalov, Jason Baldridge and David Weiss .....	328
<i>Personalized Microblog Sentiment Classification via Adversarial Cross-lingual Multi-task Learning</i> Weichao Wang, Shi Feng, Wei Gao, Daling Wang and Yifei Zhang .....	338
<i>Cross-lingual Knowledge Graph Alignment via Graph Convolutional Networks</i> Zhichun Wang, Qingsong Lv, Xiaohan Lan and Yu Zhang .....	349
<i>Cross-lingual Lexical Sememe Prediction</i> Fanchao Qi, Yankai Lin, Maosong Sun, Hao Zhu, Ruobing Xie and Zhiyuan Liu .....	358
<i>Neural Cross-Lingual Named Entity Recognition with Minimal Resources</i> Jiateng Xie, Zhilin Yang, Graham Neubig, Noah A. Smith and Jaime Carbonell .....	369
<i>A Stable and Effective Learning Strategy for Trainable Greedy Decoding</i> Yun Chen, Victor O.K. Li, Kyunghyun Cho and Samuel Bowman .....	380
<i>Addressing Troublesome Words in Neural Machine Translation</i> Yang Zhao, Jiajun Zhang, Zhongjun He, Chengqing Zong and Hua Wu .....	391

<i>Top-down Tree Structured Decoding with Syntactic Connections for Neural Machine Translation and Parsing</i>	
Jetic Gū, Hassan S. Shavarani and Anoop Sarkar . . . . .	401
<i>XL-NBT: A Cross-lingual Neural Belief Tracking Framework</i>	
Wenhu Chen, Jianshu Chen, Yu Su, Xin Wang, Dong Yu, Xifeng Yan and William Yang Wang	414
<i>Contextual Parameter Generation for Universal Neural Machine Translation</i>	
Emmanouil Antonios Platanios, Mrinmaya Sachan, Graham Neubig and Tom Mitchell . . . . .	425
<i>Back-Translation Sampling by Targeting Difficult Words in Neural Machine Translation</i>	
Marzieh Fadaee and Christof Monz . . . . .	436
<i>Multi-Domain Neural Machine Translation with Word-Level Domain Context Discrimination</i>	
Jiali Zeng, Jinsong Su, Huating Wen, Yang Liu, Jun Xie, Yongjing Yin and Jianqiang Zhao . . .	447
<i>A Discriminative Latent-Variable Model for Bilingual Lexicon Induction</i>	
Sebastian Ruder, Ryan Cotterell, Yova Kementchedjheva and Anders Søgaard . . . . .	458
<i>Non-Adversarial Unsupervised Word Translation</i>	
Yedid Hoshen and Lior Wolf . . . . .	469
<i>Semi-Autoregressive Neural Machine Translation</i>	
Chunqi Wang, Ji Zhang and Haiqing Chen . . . . .	479
<i>Understanding Back-Translation at Scale</i>	
Sergey Edunov, Myle Ott, Michael Auli and David Grangier . . . . .	489
<i>Bootstrapping Transliteration with Constrained Discovery for Low-Resource Languages</i>	
Shyam Upadhyay, Jordan Kodner and Dan Roth . . . . .	501
<i>NORMA: Neighborhood Sensitive Maps for Multilingual Word Embeddings</i>	
Ndapa Nakashole . . . . .	512
<i>Adaptive Multi-pass Decoder for Neural Machine Translation</i>	
Xinwei Geng, Xiaocheng Feng, Bing Qin and Ting Liu . . . . .	523
<i>Improving the Transformer Translation Model with Document-Level Context</i>	
Jiacheng Zhang, Huanbo Luan, Maosong Sun, Feifei Zhai, Jingfang Xu, Min Zhang and Yang Liu	533
<i>MTNT: A Testbed for Machine Translation of Noisy Text</i>	
Paul Michel and Graham Neubig . . . . .	543
<i>SimpleQuestions Nearly Solved: A New Upperbound and Baseline Approach</i>	
Michael Petrochuk and Luke Zettlemoyer . . . . .	554
<i>Phrase-Indexed Question Answering: A New Challenge for Scalable Document Comprehension</i>	
Minjoon Seo, Tom Kwiatkowski, Ankur Parikh, Ali Farhadi and Hannaneh Hajishirzi . . . . .	559
<i>Ranking Paragraphs for Improving Answer Recall in Open-Domain Question Answering</i>	
Jinhyuk Lee, Seongjun Yun, Hyunjae Kim, Miyoung Ko and Jaewoo Kang . . . . .	565
<i>Cut to the Chase: A Context Zoom-in Network for Reading Comprehension</i>	
Sathish Reddy Indurthi, Seunghak Yu, Seohyun Back and Heriberto Cuayahuitl . . . . .	570

<i>Adaptive Document Retrieval for Deep Question Answering</i> Bernhard Kratzwald and Stefan Feuerriegel .....	576
<i>Why is unsupervised alignment of English embeddings from different algorithms so hard?</i> Mareike Hartmann, Yova Kementchedjieva and Anders Søgaard .....	582
<i>Quantifying Context Overlap for Training Word Embeddings</i> Yimeng Zhuang, Jinghui Xie, Yinhe Zheng and Xuan Zhu .....	587
<i>Neural Latent Relational Analysis to Capture Lexical Semantic Relations in a Vector Space</i> Koki Washio and Tsuneaki Kato .....	594
<i>Generalizing Word Embeddings using Bag of Subwords</i> Jinman Zhao, Sidharth Mudgal and Yingyu Liang .....	601
<i>Neural Metaphor Detection in Context</i> Ge Gao, Eunsol Choi, Yejin Choi and Luke Zettlemoyer .....	607
<i>Distant Supervision from Disparate Sources for Low-Resource Part-of-Speech Tagging</i> Barbara Plank and Željko Agić .....	614
<i>Unsupervised Bilingual Lexicon Induction via Latent Variable Models</i> Zi-Yi Dou, Zhi-Hao Zhou and Shujian Huang .....	621
<i>Learning Unsupervised Word Translations Without Adversaries</i> Tanmoy Mukherjee, Makoto Yamada and Timothy Hospedales .....	627
<i>Adversarial Training for Multi-task and Multi-lingual Joint Modeling of Utterance Intent Classification</i> Ryo Masumura, Yusuke Shinohara, Ryuichiro Higashinaka and Yushi Aono .....	633
<i>Surprisingly Easy Hard-Attention for Sequence to Sequence Learning</i> Shiv Shankar, Siddhant Garg and Sunita Sarawagi .....	640
<i>Joint Learning for Emotion Classification and Emotion Cause Detection</i> Ying Chen, Wenjun Hou, Xiyao Cheng and Shoushan Li .....	646
<i>Exploring Optimism and Pessimism in Twitter Using Deep Learning</i> Cornelia Caragea, Liviu P. Dinu and Bogdan Dumitru .....	652
<i>Predicting News Headline Popularity with Syntactic and Semantic Knowledge Using Multi-Task Learning</i> Sotiris Lamprinidis, Daniel Hardt and Dirk Hovy .....	659
<i>Hybrid Neural Attention for Agreement/Disagreement Inference in Online Debates</i> Di Chen, Jiachen Du, Lidong Bing and Ruifeng Xu .....	665
<i>Increasing In-Class Similarity by Retrofitting Embeddings with Demographic Information</i> Dirk Hovy and Tommaso Fornaciari .....	671
<i>A Syntactically Constrained Bidirectional-Asynchronous Approach for Emotional Conversation Generation</i> Jingyuan Li and Xiao Sun .....	678
<i>Auto-Dialabel: Labeling Dialogue Data with Unsupervised Learning</i> Chen Shi, Qi Chen, Lei Sha, Sujian Li, Xu Sun, Houfeng Wang and Lintao Zhang .....	684

<i>Extending Neural Generative Conversational Model using External Knowledge Sources</i> Prasanna Parthasarathi and Joelle Pineau .....	690
<i>Modeling Temporality of Human Intentions by Domain Adaptation</i> Xiaolei Huang, Lixing Liu, Kate Carey, Joshua Woolley, Stefan Scherer and Brian Borsari . . . .	696
<i>An Auto-Encoder Matching Model for Learning Utterance-Level Semantic Dependency in Dialogue Generation</i> Liangchen Luo, Jingjing Xu, Junyang Lin, Qi Zeng and Xu Sun .....	702
<i>A Dataset for Document Grounded Conversations</i> Kangyan Zhou, Shrimai Prabhunoye and Alan W Black .....	708
<i>Out-of-domain Detection based on Generative Adversarial Network</i> Seonghan Ryu, Sangjun Koo, Hwanjo Yu and Gary Geunbae Lee .....	714
<i>Listening Comprehension over Argumentative Content</i> Shachar Mirkin, Guy Moshkovich, Matan Orbach, Lili Kotlerman, Yoav Kantor, Tamar Lavee, Michal Jacovi, Yonatan Bilu, Ranit Aharonov and Noam Slonim .....	719
<i>Using active learning to expand training data for implicit discourse relation recognition</i> Yang Xu, Yu Hong, Huibin Ruan, Jianmin Yao, Min Zhang and Guodong Zhou .....	725
<i>Learning To Split and Rephrase From Wikipedia Edit History</i> Jan A. Botha, Manaal Faruqui, John Alex, Jason Baldridge and Dipanjan Das .....	732
<i>BLEU is Not Suitable for the Evaluation of Text Simplification</i> Elior Sulem, Omri Abend and Ari Rappoport .....	738
<i>S2SPMN: A Simple and Effective Framework for Response Generation with Relevant Information</i> Jiaxin Pei and Chenliang Li .....	745
<i>Improving Reinforcement Learning Based Image Captioning with Natural Language Prior</i> Tszhang Guo, Shiyu Chang, Mo Yu and Kun Bai .....	751
<i>Training for Diversity in Image Paragraph Captioning</i> Luke Melas-Kyriazi, Alexander Rush and George Han .....	757
<i>A Graph-theoretic Summary Evaluation for ROUGE</i> Elaheh ShafieiBavani, Mohammad Ebrahimi, Raymond Wong and Fang Chen .....	762
<i>Guided Neural Language Generation for Abstractive Summarization using Abstract Meaning Representation</i> Hardy Hardy and Andreas Vlachos .....	768
<i>Evaluating Multiple System Summary Lengths: A Case Study</i> Ori Shapira, David Gabay, Hadar Ronen, Judit Bar-Ilan, Yael Amsterdamer, Ani Nenkova and Ido Dagan .....	774
<i>Neural Latent Extractive Document Summarization</i> Xingxing Zhang, Mirella Lapata, Furu Wei and Ming Zhou .....	779
<i>On the Abtractiveness of Neural Document Summarization</i> Fangfang Zhang, Jin-ge Yao and Rui Yan .....	785

<i>Automatic Essay Scoring Incorporating Rating Schema via Reinforcement Learning</i> Yucheng Wang, Zhongyu Wei, Yaqian Zhou and Xuanjing Huang .....	791
<i>Identifying Well-formed Natural Language Questions</i> Manaal Faruqui and Dipanjan Das .....	798
<i>Self-Governing Neural Networks for On-Device Short Text Classification</i> Sujith Ravi and Zornitsa Kozareva .....	804
<i>HFT-CNN: Learning Hierarchical Category Structure for Multi-label Short Text Categorization</i> Kazuya Shimura, Jiyi Li and Fumiyo Fukumoto .....	811
<i>A Hierarchical Neural Attention-based Text Classifier</i> Koustuv Sinha, Yue Dong, Jackie Chi Kit Cheung and Derek Ruths .....	817
<i>Labeled Anchors and a Scalable, Transparent, and Interactive Classifier</i> Jeffrey Lund, Stephen Cowley, Wilson Fearn, Emily Hales and Kevin Seppi .....	824
<i>Coherence-Aware Neural Topic Modeling</i> Ran Ding, Ramesh Nallapati and Bing Xiang .....	830
<i>Utilizing Character and Word Embeddings for Text Normalization with Sequence-to-Sequence Models</i> Daniel Watson, Nasser Zalmout and Nizar Habash .....	837
<i>Topic Intrusion for Automatic Topic Model Evaluation</i> Shraey Bhatia, Jey Han Lau and Timothy Baldwin .....	844
<i>Supervised and Unsupervised Methods for Robust Separation of Section Titles and Prose Text in Web Documents</i> Abhijith Athreya Mysore Gopinath, Shomir Wilson and Norman Sadeh .....	850
<i>SwitchOut: an Efficient Data Augmentation Algorithm for Neural Machine Translation</i> Xinyi Wang, Hieu Pham, Zihang Dai and Graham Neubig .....	856
<i>Improving Unsupervised Word-by-Word Translation with Language Model and Denoising Autoencoder</i> Yunsu Kim, Jiahui Geng and Hermann Ney .....	862
<i>Decipherment of Substitution Ciphers with Neural Language Models</i> Nishant Kambhatla, Anahita Mansouri Bigvand and Anoop Sarkar .....	869
<i>Rapid Adaptation of Neural Machine Translation to New Languages</i> Graham Neubig and Junjie Hu .....	875
<i>Compact Personalized Models for Neural Machine Translation</i> Joern Wuebker, Patrick Simianer and John DeNero .....	881
<i>Self-Governing Neural Networks for On-Device Short Text Classification</i> Sujith Ravi and Zornitsa Kozareva .....	887
<i>Supervised Domain Enablement Attention for Personalized Domain Classification</i> Joo-Kyung Kim and Young-Bum Kim .....	894
<i>A Deep Neural Network Sentence Level Classification Method with Context Information</i> Xingyi Song, Johann Petrak and Angus Roberts .....	900



<i>Towards Dynamic Computation Graphs via Sparse Latent Structure</i> Vlad Niculae, André F. T. Martins and Claire Cardie .....	905
<i>Convolutional Neural Networks with Recurrent Neural Filters</i> Yi Yang .....	912
<i>Exploiting Rich Syntactic Information for Semantic Parsing with Graph-to-Sequence Model</i> Kun Xu, Lingfei Wu, Zhiguo Wang, Mo Yu, Liwei Chen and Vadim Sheinin.....	918
<i>Retrieval-Based Neural Code Generation</i> Shirley Anugrah Hayati, Raphael Olivier, Pravalika Avvaru, Pengcheng Yin, Anthony Tomasic and Graham Neubig .....	925
<i>SQL-to-Text Generation with Graph-to-Sequence Model</i> Kun Xu, Lingfei Wu, Zhiguo Wang, Yansong Feng and Vadim Sheinin .....	931
<i>Generating Syntactic Paraphrases</i> Emilie Colin and Claire Gardent .....	937
<i>Neural-Davidsonian Semantic Proto-role Labeling</i> Rachel Rudinger, Adam Teichert, Ryan Culkin, Sheng Zhang and Benjamin Van Durme .....	944
<i>Conversational Decision-Making Model for Predicting the King’s Decision in the Annals of the Joseon Dynasty</i> JinYeong Bak and Alice Oh .....	956
<i>Toward Fast and Accurate Neural Discourse Segmentation</i> Yizhong Wang, Sujian Li and Jingfeng Yang .....	962
<i>A Dataset for Telling the Stories of Social Media Videos</i> Spandana Gella, Mike Lewis and Marcus Rohrbach .....	968
<i>Cascaded Mutual Modulation for Visual Reasoning</i> Yiqun Yao, Jiaming Xu, Feng Wang and Bo Xu .....	975
<i>How agents see things: On visual representations in an emergent language game</i> Diane Bouchacourt and Marco Baroni .....	981
<i>Attention-Based Capsule Networks with Dynamic Routing for Relation Extraction</i> Ningyu Zhang, Shumin Deng, Zhanling Sun, Xi Chen, Wei Zhang and Huajun Chen .....	986
<i>Put It Back: Entity Typing with Language Model Enhancement</i> Ji Xin, Hao Zhu, Xu Han, Zhiyuan Liu and Maosong Sun.....	993
<i>Event Detection with Neural Networks: A Rigorous Empirical Evaluation</i> Walker Orr, Prasad Tadepalli and Xiaoli Fern.....	999
<i>PubSE: A Hierarchical Model for Publication Extraction from Academic Homepages</i> Yiqing Zhang, Jianzhong Qi, Rui Zhang and Chuandong Yin .....	1005
<i>A Neural Transition-based Model for Nested Mention Recognition</i> Bailin Wang, Wei Lu, Yu Wang and Hongxia Jin.....	1011
<i>Genre Separation Network with Adversarial Training for Cross-genre Relation Extraction</i> Ge Shi, Chong Feng, Lifu Huang, Boliang Zhang, Heng Ji, Lejian Liao and Heyan Huang ...	1018

<i>Effective Use of Context in Noisy Entity Linking</i> David Mueller and Greg Durrett .....	1024
<i>Exploiting Contextual Information via Dynamic Memory Network for Event Detection</i> Shaobo Liu, Rui Cheng, Xiaoming Yu and Xueqi Cheng .....	1030
<i>Do explanations make VQA models more predictable to a human?</i> Arjun Chandrasekaran, Viraj Prabhu, Deshraj Yadav, Prithvijit Chattopadhyay and Devi Parikh	1036
<i>Facts That Matter</i> Marco Ponza, Luciano Del Corro and Gerhard Weikum .....	1043
<i>Entity Tracking Improves Cloze-style Reading Comprehension</i> Luong Hoang, Sam Wiseman and Alexander Rush .....	1049
<i>Adversarial Domain Adaptation for Duplicate Question Detection</i> Darsh Shah, Tao Lei, Alessandro Moschitti, Salvatore Romeo and Preslav Nakov .....	1056
<i>Translating a Math Word Problem to a Expression Tree</i> Lei Wang, Yan Wang, Deng Cai, Dongxiang Zhang and Xiaojiang Liu .....	1064
<i>Semantic Linking in Convolutional Neural Networks for Answer Sentence Selection</i> Massimo Nicosia and Alessandro Moschitti .....	1070
<i>A dataset and baselines for sequential open-domain question answering</i> Ahmed Elgohary, Chen Zhao and Jordan Boyd-Graber .....	1077
<i>Improving the results of string kernels in sentiment analysis and Arabic dialect identification by adapting them to your test set</i> Radu Tudor Ionescu and Andrei M. Butnaru .....	1084
<i>Parameterized Convolutional Neural Networks for Aspect Level Sentiment Classification</i> Binxuan Huang and Kathleen Carley .....	1091
<i>Improving Multi-label Emotion Classification via Sentiment Classification with Dual Attention Transfer Network</i> Jianfei Yu, Luis Marujo, Jing Jiang, Pradeep Karuturi and William Brendel .....	1097
<i>Learning Sentiment Memories for Sentiment Modification without Parallel Data</i> Yi Zhang, Jingjing Xu, Pengcheng Yang and Xu Sun .....	1103
<i>Joint Aspect and Polarity Classification for Aspect-based Sentiment Analysis with End-to-End Neural Networks</i> Martin Schmitt, Simon Steinheber, Konrad Schreiber and Benjamin Roth .....	1109
<i>Representing Social Media Users for Sarcasm Detection</i> Y. Alex Kolchinski and Christopher Potts .....	1115
<i>Syntactical Analysis of the Weaknesses of Sentiment Analyzers</i> Rohil Verma, Samuel Kim and David Walter .....	1122
<i>Is Nike female? Exploring the role of sound symbolism in predicting brand name gender</i> Sridhar Moorthy, Ruth Pogacar, Samin Khan and Yang Xu .....	1128
<i>Improving Large-Scale Fact-Checking using Decomposable Attention Models and Lexical Tagging</i> Nayeon Lee, Chien-Sheng Wu and Pascale Fung .....	1133

<i>Harnessing Popularity in Social Media for Extractive Summarization of Online Conversations</i> Ryuji Kano, Yasuhide Miura, Motoki Taniguchi, Yan-Ying Chen, Francine Chen and Tomoko Ohkuma .....	1139
<i>Identifying Locus of Control in Social Media Language</i> Masoud Rouhizadeh, Kokil Jaidka, Laura Smith, H. Andrew Schwartz, Anneke Buffone and Lyle Ungar .....	1146
<i>Somm: Into the Model</i> Shengli Hu .....	1153
<i>Fine-Grained Emotion Detection in Health-Related Online Posts</i> Hamed Khanpour and Cornelia Caragea .....	1160
<i>The Remarkable Benefit of User-Level Aggregation for Lexical-based Population-Level Predictions</i> Salvatore Giorgi, Daniel Preoțiu-Pietro, Anneke Buffone, Daniel Rieman, Lyle Ungar and H. Andrew Schwartz .....	1167
<i>Deterministic Non-Autoregressive Neural Sequence Modeling by Iterative Refinement</i> Jason Lee, Elman Mansimov and Kyunghyun Cho .....	1173
<i>Large Margin Neural Language Model</i> Jiaji Huang, Yi Li, Wei Ping and Liang Huang .....	1183
<i>Targeted Syntactic Evaluation of Language Models</i> Rebecca Marvin and Tal Linzen .....	1192
<i>Rational Recurrences</i> Hao Peng, Roy Schwartz, Sam Thomson and Noah A. Smith .....	1203
<i>Efficient Contextualized Representation: Language Model Pruning for Sequence Labeling</i> Liyuan Liu, Xiang Ren, Jingbo Shang, Xiaotao Gu, Jian Peng and Jiawei Han .....	1215
<i>Automatic Event Salience Identification</i> Zhengzhong Liu, Chenyan Xiong, Teruko Mitamura and Eduard Hovy .....	1226
<i>Temporal Information Extraction by Predicting Relative Time-lines</i> Artuur Leeuwenberg and Marie-Francine Moens .....	1237
<i>Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation</i> Xiao Liu, Zhunchen Luo and Heyan Huang .....	1247
<i>RESIDE: Improving Distantly-Supervised Neural Relation Extraction using Side Information</i> Shikhar Vashishth, Rishabh Joshi, Sai Suman Prayaga, Chiranjib Bhattacharyya and Partha Talukdar	1257
<i>Collective Event Detection via a Hierarchical and Bias Tagging Networks with Gated Multi-level Attention Mechanisms</i> Yubo Chen, Hang Yang, Kang Liu, Jun Zhao and Yantao Jia .....	1267
<i>Valency-Augmented Dependency Parsing</i> Tianze Shi and Lillian Lee .....	1277
<i>Unsupervised Learning of Syntactic Structure with Invertible Neural Projections</i> Junxian He, Graham Neubig and Taylor Berg-Kirkpatrick .....	1292

<i>Dynamic Oracles for Top-Down and In-Order Shift-Reduce Constituent Parsing</i> Daniel Fernández-González and Carlos Gómez-Rodríguez .....	1303
<i>Constituent Parsing as Sequence Labeling</i> Carlos Gómez-Rodríguez and David Vilares .....	1314
<i>Synthetic Data Made to Order: The Case of Parsing</i> Dingquan Wang and Jason Eisner .....	1325
<i>Tell-and-Answer: Towards Explainable Visual Question Answering using Attributes and Captions</i> Qing Li, Jianlong Fu, Dongfei Yu, Tao Mei and Jiebo Luo .....	1338
<i>Learning a Policy for Opportunistic Active Learning</i> Aishwarya Padmakumar, Peter Stone and Raymond Mooney .....	1347
<i>RecipeQA: A Challenge Dataset for Multimodal Comprehension of Cooking Recipes</i> Semih Yagcioglu, Aykut Erdem, Erkut Erdem and Nazli Ikizler-Cinbis .....	1358
<i>TVQA: Localized, Compositional Video Question Answering</i> Jie Lei, Licheng Yu, Mohit Bansal and Tamara Berg .....	1369
<i>Localizing Moments in Video with Temporal Language</i> Lisa Anne Hendricks, Oliver Wang, Eli Shechtman, Josef Sivic, Trevor Darrell and Bryan Russell	1380
<i>Card-660: Cambridge Rare Word Dataset - A Reliable Evaluation Framework for Rare Word Representation Models</i> Mohammad Taher Pilehvar, Dimitri Kartsaklis, Victor Prokhorov and Nigel Collier .....	1391
<i>Leveraging Gloss Knowledge in Neural Word Sense Disambiguation by Hierarchical Co-Attention</i> Fuli Luo, Tianyu Liu, Zexue He, Qiaolin Xia, Zhifang Sui and Baobao Chang .....	1402
<i>Weeding out Conventionalized Metaphors: A Corpus of Novel Metaphor Annotations</i> Erik-Lân Do Dinh, Hannah Wieland and Iryna Gurevych .....	1412
<i>Streaming word similarity mining on the cheap</i> Olof Görnerup and Daniel Gillblad .....	1425
<i>Memory, Show the Way: Memory Based Few Shot Word Representation Learning</i> Jingyuan Sun, Shaonan Wang and Chengqing Zong .....	1435
<i>Disambiguated skip-gram model</i> Karol Grzegorzczak and Marcin Kurdział .....	1445
<i>Picking Apart Story Salads</i> Su Wang, Eric Holgate, Greg Durrett and Katrin Erk .....	1455
<i>Dynamic Meta-Embeddings for Improved Sentence Representations</i> Douwe Kiela, Changhan Wang and Kyunghyun Cho .....	1466
<i>A Probabilistic Model for Joint Learning of Word Embeddings from Texts and Images</i> Melissa Ailem, Bowen Zhang, Aurélien Bellet, Pascal Denis and Fei Sha .....	1478
<i>Transfer and Multi-Task Learning for Noun–Noun Compound Interpretation</i> Murhaf Fares, Stephan Oepen and Erik Velldal .....	1488

<i>Dissecting Contextual Word Embeddings: Architecture and Representation</i> Matthew Peters, Mark Neumann, Luke Zettlemoyer and Wen-tau Yih .....	1499
<i>Preposition Sense Disambiguation and Representation</i> Hongyu Gong, Jiaqi Mu, Suma Bhat and Pramod Viswanath .....	1510
<i>Auto-Encoding Dictionary Definitions into Consistent Word Embeddings</i> Tom Bosc and Pascal Vincent .....	1522
<i>Spot the Odd Man Out: Exploring the Associative Power of Lexical Resources</i> Gabriel Stanovsky and Mark Hopkins .....	1533
<i>Neural Multitask Learning for Simile Recognition</i> Lizhen Liu, Xiao Hu, Wei Song, Ruiji Fu, Ting Liu and Guoping Hu .....	1543
<i>Structured Alignment Networks for Matching Sentences</i> Yang Liu, Matt Gardner and Mirella Lapata .....	1554
<i>Compare, Compress and Propagate: Enhancing Neural Architectures with Alignment Factorization for Natural Language Inference</i> Yi Tay, Anh Tuan Luu and Siu Cheung Hui .....	1565
<i>Convolutional Interaction Network for Natural Language Inference</i> Jingjing Gong, Xipeng Qiu, Xinchu Chen, Dong Liang and Xuanjing Huang .....	1576
<i>Lessons from Natural Language Inference in the Clinical Domain</i> Alexey Romanov and Chaitanya Shivade .....	1586
<i>Question Generation from SQL Queries Improves Neural Semantic Parsing</i> Daya Guo, Yibo Sun, Duyu Tang, Nan Duan, Jian Yin, Hong Chi, James Cao, Peng Chen and Ming Zhou .....	1597
<i>SemRegex: A Semantics-Based Approach for Generating Regular Expressions from Natural Language Specifications</i> Zexuan Zhong, Jiaqi Guo, Wei Yang, Jian Peng, Tao Xie, Jian-Guang Lou, Ting Liu and Dongmei Zhang .....	1608
<i>Decoupling Structure and Lexicon for Zero-Shot Semantic Parsing</i> Jonathan Herzig and Jonathan Berant .....	1619
<i>A Span Selection Model for Semantic Role Labeling</i> Hiroki Ouchi, Hiroyuki Shindo and Yuji Matsumoto .....	1630
<i>Mapping Language to Code in Programmatic Context</i> Srinivasan Iyer, Ioannis Konstas, Alvin Cheung and Luke Zettlemoyer .....	1643
<i>SyntaxSQLNet: Syntax Tree Networks for Complex and Cross-Domain Text-to-SQL Task</i> Tao Yu, Michihiro Yasunaga, Kai Yang, Rui Zhang, Dongxu Wang, Zifan Li and Dragomir Radev .....	1653
<i>Cross-lingual Compositional Semantic Parsing</i> Sheng Zhang, Xutai Ma, Rachel Rudinger, Kevin Duh and Benjamin Van Durme .....	1664
<i>Learning to Learn Semantic Parsers from Natural Language Supervision</i> Igor Labutov, Bishan Yang and Tom Mitchell .....	1676

<i>DeepCx: A transition-based approach for shallow semantic parsing with complex constructional triggers</i> Jesse Dunietz, Jaime Carbonell and Lori Levin .....	1691
<i>What It Takes to Achieve 100% Condition Accuracy on WikiSQL</i> Semih Yavuz, Izzeddin Gur, Yu Su and Xifeng Yan .....	1702
<i>Better Transition-Based AMR Parsing with a Refined Search Space</i> Zhijiang Guo and Wei Lu .....	1712
<i>Heuristically Informed Unsupervised Idiom Usage Recognition</i> Changsheng Liu and Rebecca Hwa .....	1723
<i>Coming to Your Senses: on Controls and Evaluation Sets in Polysemy Research</i> Haim Dubossarsky, Eitan Grossman and Daphna Weinshall .....	1732
<i>Predicting Semantic Relations using Global Graph Properties</i> Yuval Pinter and Jacob Eisenstein .....	1741
<i>Learning Scalar Adjective Intensity from Paraphrases</i> Anne Cocos, Veronica Wharton, Ellie Pavlick, Marianna Apidianaki and Chris Callison-Burch .....	1752
<i>Pointwise HSIC: A Linear-Time Kernelized Co-occurrence Norm for Sparse Linguistic Expressions</i> Sho Yokoi, Sosuke Kobayashi, Kenji Fukumizu, Jun Suzuki and Kentaro Inui .....	1763
<i>Neural Related Work Summarization with a Joint Context-driven Attention Mechanism</i> Yongzhen Wang, Xiaozhong Liu and Zheng Gao .....	1776
<i>Improving Neural Abstractive Document Summarization with Explicit Information Selection Modeling</i> Wei Li, Xinyan Xiao, Yajuan Lyu and Yuanzhuo Wang .....	1787
<i>Don't Give Me the Details, Just the Summary! Topic-Aware Convolutional Neural Networks for Extreme Summarization</i> Shashi Narayan, Shay B. Cohen and Mirella Lapata .....	1797
<i>Improving Abstraction in Text Summarization</i> Wojciech Kryściński, Romain Paulus, Caiming Xiong and Richard Socher .....	1808
<i>Content Selection in Deep Learning Models of Summarization</i> Chris Kedzie, Kathleen McKeown and Hal Daume III .....	1818
<i>Improved Semantic-Aware Network Embedding with Fine-Grained Word Alignment</i> Dinghan Shen, Xinyuan Zhang, Ricardo Henao and Lawrence Carin .....	1829
<i>Learning Context-Sensitive Convolutional Filters for Text Processing</i> Dinghan Shen, Martin Renqiang Min, Yitong Li and Lawrence Carin .....	1839
<i>Deep Relevance Ranking Using Enhanced Document-Query Interactions</i> Ryan McDonald, George Brokos and Ion Androutsopoulos .....	1849
<i>Learning Neural Representation for CLIR with Adversarial Framework</i> Bo Li and Ping Cheng .....	1861
<i>AD3: Attentive Deep Document Dater</i> Swayambhu Nath Ray, Shib Sankar Dasgupta and Partha Talukdar .....	1871

<i>Gromov-Wasserstein Alignment of Word Embedding Spaces</i>	
David Alvarez-Melis and Tommi Jaakkola .....	1881
<i>Deep Probabilistic Logic: A Unifying Framework for Indirect Supervision</i>	
Hai Wang and Hoifung Poon .....	1891
<i>Deriving Machine Attention from Human Rationales</i>	
Yujia Bao, Shiyu Chang, Mo Yu and Regina Barzilay .....	1903
<i>Semi-Supervised Sequence Modeling with Cross-View Training</i>	
Kevin Clark, Minh-Thang Luong, Christopher D. Manning and Quoc Le .....	1914
<i>A Probabilistic Annotation Model for Crowdsourcing Coreference</i>	
Silviu Paun, Jon Chamberlain, Udo Kruschwitz, Juntao Yu and Massimo Poesio .....	1926
<i>A Deterministic Algorithm for Bridging Anaphora Resolution</i>	
Yufang Hou .....	1938
<i>A Knowledge Hunting Framework for Common Sense Reasoning</i>	
Ali Emami, Noelia De La Cruz, Adam Trischler, Kaheer Suleman and Jackie Chi Kit Cheung .....	1949
<i>Mapping Text to Knowledge Graph Entities using Multi-Sense LSTMs</i>	
Dimitri Kartsaklis, Mohammad Taher Pilehvar and Nigel Collier .....	1959
<i>Differentiating Concepts and Instances for Knowledge Graph Embedding</i>	
Xin Lv, Lei Hou, Juanzi Li and Zhiyuan Liu .....	1971
<i>One-Shot Relational Learning for Knowledge Graphs</i>	
Wenhan Xiong, Mo Yu, Shiyu Chang, Xiaoxiao Guo and William Yang Wang .....	1980
<i>Regular Expression Guided Entity Mention Mining from Noisy Web Data</i>	
Shanshan Zhang, Lihong He, Slobodan Vucetic and Eduard Dragut .....	1991
<i>HyTE: Hyperplane-based Temporally aware Knowledge Graph Embedding</i>	
Shib Sankar Dasgupta, Swayambhu Nath Ray and Partha Talukdar .....	2001
<i>Neural Adaptation Layers for Cross-domain Named Entity Recognition</i>	
Bill Yuchen Lin and Wei Lu .....	2012
<i>Entity Linking within a Social Media Platform: A Case Study on Yelp</i>	
Hongliang Dai, Yangqiu Song, Liwei Qiu and Rijia Liu .....	2023
<i>Annotation of a Large Clinical Entity Corpus</i>	
Pinal Patel, Disha Davey, Vishal Panchal and Parth Pathak .....	2033
<i>Visual Supervision in Bootstrapped Information Extraction</i>	
Matthew Berger, Ajay Nagesh, Joshua Levine, Mihai Surdeanu and Helen Zhang .....	2043
<i>Learning Named Entity Tagger using Domain-Specific Dictionary</i>	
Jingbo Shang, Liyuan Liu, Xiaotao Gu, Xiang Ren, Teng Ren and Jiawei Han .....	2054
<i>Zero-Shot Open Entity Typing as Type-Compatible Grounding</i>	
Ben Zhou, Daniel Khashabi, Chen-Tse Tsai and Dan Roth .....	2065
<i>Attention-Guided Answer Distillation for Machine Reading Comprehension</i>	
Minghao Hu, Yuxing Peng, Furu Wei, Zhen Huang, Dongsheng Li, Nan Yang and Ming Zhou .....	2077

<i>Interpretation of Natural Language Rules in Conversational Machine Reading</i> Marzieh Saeidi, Max Bartolo, Patrick Lewis, Sameer Singh, Tim Rocktäschel, Mike Sheldon, Guillaume Bouchard and Sebastian Riedel .....	2087
<i>A State-transition Framework to Answer Complex Questions over Knowledge Base</i> Sen Hu, Lei Zou and Xinbo Zhang .....	2098
<i>A Multi-answer Multi-task Framework for Real-world Machine Reading Comprehension</i> Jiahua Liu, Wan Wei, Maosong Sun, Hao Chen, Yantao Du and Dekang Lin .....	2109
<i>Logician and Orator: Learning from the Duality between Language and Knowledge in Open Domain</i> Mingming Sun, Xu Li and Ping Li .....	2119
<i>MemoReader: Large-Scale Reading Comprehension through Neural Memory Controller</i> Seohyun Back, Seunghak Yu, Sathish Reddy Indurthi, Jihie Kim and Jaegul Choo .....	2131
<i>Multi-Granular Sequence Encoding via Dilated Compositional Units for Reading Comprehension</i> Yi Tay, Anh Tuan Luu and Siu Cheung Hui .....	2141
<i>Neural Compositional Denotational Semantics for Question Answering</i> Nitish Gupta and Mike Lewis .....	2152
<i>Cross-Pair Text Representations for Answer Sentence Selection</i> Kateryna Tymoshenko and Alessandro Moschitti .....	2162
<i>QuAC: Question Answering in Context</i> Eunsol Choi, He He, Mohit Iyyer, Mark Yatskar, Wen-tau Yih, Yejin Choi, Percy Liang and Luke Zettlemoyer .....	2174
<i>Knowledge Base Question Answering via Encoding of Complex Query Graphs</i> Kangqi Luo, Fengli Lin, Xusheng Luo and Kenny Zhu .....	2185
<i>Neural Relation Extraction via Inner-Sentence Noise Reduction and Transfer Learning</i> Tianyi Liu, Xinsong Zhang, Wanhao Zhou and Weijia Jia .....	2195
<i>Graph Convolution over Pruned Dependency Trees Improves Relation Extraction</i> Yuhao Zhang, Peng Qi and Christopher D. Manning .....	2205
<i>Multi-Level Structured Self-Attentions for Distantly Supervised Relation Extraction</i> Jinhua Du, Jingguang Han, Andy Way and Dadong Wan .....	2216
<i>N-ary Relation Extraction using Graph-State LSTM</i> Linfeng Song, Yue Zhang, Zhiguo Wang and Daniel Gildea .....	2226
<i>Hierarchical Relation Extraction with Coarse-to-Fine Grained Attention</i> Xu Han, Pengfei Yu, Zhiyuan Liu, Maosong Sun and Peng Li .....	2236
<i>Label-Free Distant Supervision for Relation Extraction via Knowledge Graph Embedding</i> Guanying Wang, Wen Zhang, Ruoxu Wang, Yalin Zhou, Xi Chen, Wei Zhang, Hai Zhu and Huajun Chen .....	2246
<i>Extracting Entities and Relations with Joint Minimum Risk Training</i> Changzhi Sun, Yuanbin Wu, Man Lan, Shiliang Sun, Wenting Wang, Kuang-Chih Lee and Kewen Wu .....	2256



<i>Large-scale Exploration of Neural Relation Classification Architectures</i>	
Hoang-Quynh Le, Duy-Cat Can, Sinh T. Vu, Thanh Hai Dang, Mohammad Taher Pilehvar and Nigel Collier . . . . .	2266
<i>Possessors Change Over Time: A Case Study with Artworks</i>	
Dhivya Chinnappa and Eduardo Blanco . . . . .	2278
<i>Using Lexical Alignment and Referring Ability to Address Data Sparsity in Situated Dialog Reference Resolution</i>	
Todd Shore and Gabriel Skantze . . . . .	2288
<i>Subgoal Discovery for Hierarchical Dialogue Policy Learning</i>	
Da Tang, Xiujun Li, Jianfeng Gao, Chong Wang, Lihong Li and Tony Jebara . . . . .	2298
<i>Supervised Clustering of Questions into Intents for Dialog System Applications</i>	
Iryna Haponchyk, Antonio Uva, Seunghak Yu, Olga Uryupina and Alessandro Moschitti . . . . .	2310
<i>Towards Exploiting Background Knowledge for Building Conversation Systems</i>	
Nikita Moghe, Siddhartha Arora, Suman Banerjee and Mitesh M. Khapra . . . . .	2322
<i>Decoupling Strategy and Generation in Negotiation Dialogues</i>	
He He, Derek Chen, Anusha Balakrishnan and Percy Liang . . . . .	2333
<i>Large-scale Cloze Test Dataset Created by Teachers</i>	
Qizhe Xie, Guokun Lai, Zihang Dai and Eduard Hovy . . . . .	2344
<i>emrQA: A Large Corpus for Question Answering on Electronic Medical Records</i>	
Anusri Pampari, Preethi Raghavan, Jennifer Liang and Jian Peng . . . . .	2357
<i>HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering</i>	
Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William Cohen, Ruslan Salakhutdinov and Christopher D. Manning . . . . .	2369
<i>Can a Suit of Armor Conduct Electricity? A New Dataset for Open Book Question Answering</i>	
Todor Mihaylov, Peter Clark, Tushar Khot and Ashish Sabharwal . . . . .	2381
<i>Evaluating Theory of Mind in Question Answering</i>	
Aida Nematzadeh, Kaylee Burns, Erin Grant, Alison Gopnik and Tom Griffiths . . . . .	2392
<i>A Unified Syntax-aware Framework for Semantic Role Labeling</i>	
Zuchao Li, Shexia He, Jiaxun Cai, Zhuosheng Zhang, Hai Zhao, Gongshen Liu, Linlin Li and Luo Si . . . . .	2401
<i>Semantics as a Foreign Language</i>	
Gabriel Stanovsky and Ido Dagan . . . . .	2412
<i>An AMR Aligner Tuned by Transition-based Parser</i>	
Yijia Liu, Wanxiang Che, Bo Zheng, Bing Qin and Ting Liu . . . . .	2422
<i>Dependency-based Hybrid Trees for Semantic Parsing</i>	
Zhanming Jie and Wei Lu . . . . .	2431
<i>Policy Shaping and Generalized Update Equations for Semantic Parsing from Denotations</i>	
Dipendra Misra, Ming-Wei Chang, Xiaodong He and Wen-tau Yih . . . . .	2442

<i>Sentence Compression for Arbitrary Languages via Multilingual Pivoting</i> Jonathan Mallinson, Rico Sennrich and Mirella Lapata .....	2453
<i>Unsupervised Cross-lingual Transfer of Word Embedding Spaces</i> Ruochen Xu, Yiming Yang, Naoki Otani and Yuexin Wu .....	2465
<i>XNLI: Evaluating Cross-lingual Sentence Representations</i> Alexis Conneau, Ruty Rinott, Guillaume Lample, Adina Williams, Samuel Bowman, Holger Schwenk and Veselin Stoyanov .....	2475
<i>Joint Multilingual Supervision for Cross-lingual Entity Linking</i> Shyam Upadhyay, Nitish Gupta and Dan Roth .....	2486
<i>Fine-grained Coordinated Cross-lingual Text Stream Alignment for Endless Language Knowledge Acquisition</i> Tao Ge, Qing Dou, Heng Ji, Lei Cui, Baobao Chang, Zhifang Sui, Furu Wei and Ming Zhou ..	2496
<i>WECA: A WordNet-Encoded Collocation-Attention Network for Homographic Pun Recognition</i> Yufeng Diao, Hongfei Lin, Di Wu, Liang Yang, Kan Xu, Zhihao Yang, Jian Wang, Shaowu Zhang, Bo Xu and Dongyu Zhang .....	2507
<i>A Hybrid Approach to Automatic Corpus Generation for Chinese Spelling Check</i> Dingmin Wang, Yan Song, Jing Li, Jialong Han and Haisong Zhang .....	2517
<i>Neural Quality Estimation of Grammatical Error Correction</i> Shamil Chollampatt and Hwee Tou Ng .....	2528
<i>Transferring from Formal Newswire Domain with Hypernet for Twitter POS Tagging</i> Tao Gui, Qi Zhang, Jingjing Gong, Minlong Peng, di liang, Keyu Ding and Xuanjing Huang ..	2540
<i>Free as in Free Word Order: An Energy Based Model for Word Segmentation and Morphological Tagging in Sanskrit</i> Amrith Krishna, Bishal Santra, Sasi Prasanth Bandaru, Gaurav Sahu, Vishnu Dutt Sharma, Pavan Kumar Satuluri and Pawan Goyal .....	2550
<i>A Challenge Set and Methods for Noun-Verb Ambiguity</i> Ali Elkahky, Kellie Webster, Daniel Andor and Emily Pitler .....	2562
<i>What do character-level models learn about morphology? The case of dependency parsing</i> Clara Vania, Andreas Grivas and Adam Lopez .....	2573
<i>Learning Better Internal Structure of Words for Sequence Labeling</i> Yingwei Xin, Ethan Hart, Vibhuti Mahajan and Jean David Ruvini .....	2584
<i>ICON: Interactive Conversational Memory Network for Multimodal Emotion Detection</i> Devamanyu Hazarika, Soujanya Poria, Rada Mihalcea, Erik Cambria and Roger Zimmermann	2594
<i>Discriminative Learning of Open-Vocabulary Object Retrieval and Localization by Negative Phrase Augmentation</i> Ryota Hinami and Shin'ichi Satoh .....	2605
<i>Grounding Semantic Roles in Images</i> Carina Silberer and Manfred Pinkal .....	2616
<i>Commonsense Justification for Action Explanation</i> Shaohua Yang, Qiaozi Gao, Sari Sadiya and Joyce Chai .....	2627

<i>Learning Personas from Dialogue with Attentive Memory Networks</i> Eric Chu, Prashanth Vijayaraghavan and Deb Roy .....	2638
<i>Grounding language acquisition by training semantic parsers using captioned videos</i> Candace Ross, Andrei Barbu, Yevgeni Berzak, Battushig Myanganbayar and Boris Katz.....	2647
<i>Translating Navigation Instructions in Natural Language to a High-Level Plan for Behavioral Robot Navigation</i> Xiaoxue Zang, Ashwini Pokle, Marynel Vázquez, Kevin Chen, Juan Carlos Niebles, Alvaro Soto and Silvio Savarese.....	2657
<i>Mapping Instructions to Actions in 3D Environments with Visual Goal Prediction</i> Dipendra Misra, Andrew Bennett, Valts Blukis, Eyvind Niklasson, Max Shatkhin and Yoav Artzi	2667
<i>Deconvolutional Time Series Regression: A Technique for Modeling Temporally Diffuse Effects</i> Cory Shain and William Schuler .....	2679
<i>Is this Sentence Difficult? Do you Agree?</i> Dominique Brunato, Lorenzo De Mattei, Felice Dell’Orletta, Benedetta Iavarone and Giulia Venturi	2690
<i>Neural Transition Based Parsing of Web Queries: An Entity Based Approach</i> Rivka Malca and Roi Reichart .....	2700
<i>An Investigation of the Interactions Between Pre-Trained Word Embeddings, Character Models and POS Tags in Dependency Parsing</i> Aaron Smith, Miryam de Lhoneux, Sara Stymne and Joakim Nivre .....	2711
<i>Depth-bounding is effective: Improvements and evaluation of unsupervised PCFG induction</i> Lifeng Jin, Finale Doshi-Velez, Timothy Miller, William Schuler and Lane Schwartz .....	2721
<i>Incremental Computation of Infix Probabilities for Probabilistic Finite Automata</i> Marco Cognetta, Yo-Sub Han and Soon Chan Kwon .....	2732
<i>Syntax Encoding with Application in Authorship Attribution</i> Richong Zhang, Zhiyuan Hu, Hongyu Guo and Yongyi Mao .....	2742
<i>Sanskrit Word Segmentation Using Character-level Recurrent and Convolutional Neural Networks</i> Oliver Hellwig and Sebastian Nehrlich .....	2754
<i>Session-level Language Modeling for Conversational Speech</i> Wayne Xiong, Lingfeng Wu, Jun Zhang and Andreas Stolcke .....	2764
<i>Towards Less Generic Responses in Neural Conversation Models: A Statistical Re-weighting Method</i> Yahui Liu, Wei Bi, Jun Gao, Xiaojiang Liu, Jian Yao and Shuming Shi .....	2769
<i>Training Millions of Personalized Dialogue Agents</i> Pierre-Emmanuel Mazare, Samuel Humeau, Martin Raison and Antoine Bordes .....	2775
<i>Towards Universal Dialogue State Tracking</i> Liliang Ren, Kaige Xie, Lu Chen and Kai Yu.....	2780
<i>Semantic Parsing for Task Oriented Dialog using Hierarchical Representations</i> Sonal Gupta, Rushin Shah, Mrinal Mohit, Anuj Kumar and Mike Lewis.....	2787

<i>The glass ceiling in NLP</i> Natalie Schluter .....	2793
<i>Reducing Gender Bias in Abusive Language Detection</i> Ji Ho Park, Jamin Shin and Pascale Fung .....	2799
<i>SafeCity: Understanding Diverse Forms of Sexual Harassment Personal Stories</i> Sweta Karlekar and Mohit Bansal .....	2805
<i>Learning multiview embeddings for assessing dementia</i> Chloé Pou-Prom and Frank Rudzicz .....	2812
<i>WikiConv: A Corpus of the Complete Conversational History of a Large Online Collaborative Community</i> Yiqing Hua, Cristian Danescu-Niculescu-Mizil, Dario Taraborelli, Nithum Thain, Jeffery Sorensen and Lucas Dixon .....	2818
<i>Marginal Likelihood Training of BiLSTM-CRF for Biomedical Named Entity Recognition from Disjoint Label Sets</i> Nathan Greenberg, Trapit Bansal, Patrick Verga and Andrew McCallum .....	2824
<i>Adversarial training for multi-context joint entity and relation extraction</i> Giannis Bekoulis, Johannes Deleu, Thomas Demeester and Chris Develder .....	2830
<i>Structured Multi-Label Biomedical Text Tagging via Attentive Neural Tree Decoding</i> Gaurav Singh, James Thomas, Iain Marshall, John Shawe-Taylor and Byron C. Wallace .....	2837
<i>Deep Exhaustive Model for Nested Named Entity Recognition</i> Mohammad Golam Sohrab and Makoto Miwa .....	2843
<i>Evaluating the Utility of Hand-crafted Features in Sequence Labelling</i> Minghao Wu, Fei Liu and Trevor Cohn .....	2850
<i>Improved Dependency Parsing using Implicit Word Connections Learned from Unlabeled Data</i> Wenhui Wang, Baobao Chang and Mairgup Mansur .....	2857
<i>A Framework for Understanding the Role of Morphology in Universal Dependency Parsing</i> Mathieu Dehouck and Pascal Denis .....	2864
<i>The Lazy Encoder: A Fine-Grained Analysis of the Role of Morphology in Neural Machine Translation</i> Arianna Bisazza and Clara Tump .....	2871
<i>Imitation Learning for Neural Morphological String Transduction</i> Peter Makarov and Simon Clematide .....	2877
<i>An Encoder-Decoder Approach to the Paradigm Cell Filling Problem</i> Miikka Silfverberg and Mans Hulden .....	2883
<i>Generating Natural Language Adversarial Examples</i> Moustafa Alzantot, Yash Sharma, Ahmed Elgohary, Bo-Jhang Ho, Mani Srivastava and Kai-Wei Chang .....	2890
<i>Multi-Head Attention with Disagreement Regularization</i> Jian Li, Zhaopeng Tu, Baosong Yang, Michael R. Lyu and Tong Zhang .....	2897

<i>Deep Bayesian Active Learning for Natural Language Processing: Results of a Large-Scale Empirical Study</i>	
Aditya Siddhant and Zachary C. Lipton . . . . .	2904
<i>Bayesian Compression for Natural Language Processing</i>	
Nadezhda Chirkova, Ekaterina Lobacheva and Dmitry Vetrov . . . . .	2910
<i>Multimodal neural pronunciation modeling for spoken languages with logographic origin</i>	
Minh Nguyen, Gia H Ngo and Nancy Chen . . . . .	2916
<i>Chinese Pinyin Aided IME, Input What You Have Not Keystroked Yet</i>	
Yafang Huang and Hai Zhao . . . . .	2923
<i>Estimating Marginal Probabilities of n-grams for Recurrent Neural Language Models</i>	
Thanapon Noraset, Doug Downey and Lidong Bing . . . . .	2930
<i>How to represent a word and predict it, too: Improving tied architectures for language modelling</i>	
Kristina Gulordava, Laura Aina and Gemma Boleda . . . . .	2936
<i>The Importance of Generation Order in Language Modeling</i>	
Nicolas Ford, Daniel Duckworth, Mohammad Norouzi and George Dahl . . . . .	2942
<i>Document-Level Neural Machine Translation with Hierarchical Attention Networks</i>	
Lesly Miculicich, Dhananjay Ram, Nikolaos Pappas and James Henderson . . . . .	2947
<i>Three Strategies to Improve One-to-Many Multilingual Translation</i>	
Yining Wang, Jiajun Zhang, Feifei Zhai, Jingfang Xu and Chengqing Zong . . . . .	2955
<i>Multi-Source Syntactic Neural Machine Translation</i>	
Anna Currey and Kenneth Heafield . . . . .	2961
<i>Fixing Translation Divergences in Parallel Corpora for Neural MT</i>	
Minh Quang Pham, Josep Crego, Jean Senellart and François Yvon . . . . .	2967
<i>Adversarial Evaluation of Multimodal Machine Translation</i>	
Desmond Elliott . . . . .	2974
<i>Loss in Translation: Learning Bilingual Word Mapping with a Retrieval Criterion</i>	
Armand Joulin, Piotr Bojanowski, Tomas Mikolov, Hervé Jégou and Edouard Grave . . . . .	2979
<i>Learning When to Concentrate or Divert Attention: Self-Adaptive Attention Temperature for Neural Machine Translation</i>	
Junyang Lin, Xu Sun, Xuancheng Ren, Muyu Li and Qi Su . . . . .	2985
<i>Accelerating Asynchronous Stochastic Gradient Descent for Neural Machine Translation</i>	
Nikolay Bogoychev, Kenneth Heafield, Alham Fikri Aji and Marcin Junczys-Dowmunt . . . . .	2991
<i>Learning to Jointly Translate and Predict Dropped Pronouns with a Shared Reconstruction Mechanism</i>	
Longyue Wang, Zhaopeng Tu, Andy Way and Qun Liu . . . . .	2997
<i>Getting Gender Right in Neural Machine Translation</i>	
Eva Vanmassenhove, Christian Hardmeier and Andy Way . . . . .	3003
<i>Towards Two-Dimensional Sequence to Sequence Model in Neural Machine Translation</i>	
Parnia Bahar, Christopher Brix and Hermann Ney . . . . .	3009

<i>End-to-End Non-Autoregressive Neural Machine Translation with Connectionist Temporal Classification</i> Jindřich Libovický and Jindřich Helcl .....	3016
<i>Prediction Improves Simultaneous Neural Machine Translation</i> Ashkan Alinejad, Maryam Siahbani and Anoop Sarkar .....	3022
<i>Training Deeper Neural Machine Translation Models with Transparent Attention</i> Ankur Bapna, Mia Chen, Orhan Firat, Yuan Cao and Yonghui Wu .....	3028
<i>Context and Copying in Neural Machine Translation</i> Rebecca Knowles and Philipp Koehn .....	3034
<i>Encoding Gated Translation Memory into Neural Machine Translation</i> Qian Cao and Deyi Xiong .....	3042
<i>Automatic Post-Editing of Machine Translation: A Neural Programmer-Interpreter Approach</i> Thuy-Trang Vu and Gholamreza Haffari .....	3048
<i>Breaking the Beam Search Curse: A Study of (Re-)Scoring Methods and Stopping Criteria for Neural Machine Translation</i> Yilin Yang, Liang Huang and Mingbo Ma .....	3054
<i>Multi-Multi-View Learning: Multilingual and Multi-Representation Entity Typing</i> Yadollah Yaghoobzadeh and Hinrich Schütze .....	3060
<i>Word Embeddings for Code-Mixed Language Processing</i> Adithya Pratapa, Monojit Choudhury and Sunayana Sitaram .....	3067
<i>On the Strength of Character Language Models for Multilingual Named Entity Recognition</i> Xiaodong Yu, Stephen Mayhew, Mark Sammons and Dan Roth .....	3073
<i>Code-switched Language Models Using Dual RNNs and Same-Source Pretraining</i> Saurabh Garg, Tanmay Parekh and Preethi Jyothi .....	3078
<i>Part-of-Speech Tagging for Code-Switched, Transliterated Texts without Explicit Language Identification</i> Kelsey Ball and Dan Garrette .....	3084
<i>Zero-shot User Intent Detection via Capsule Neural Networks</i> Congying Xia, Chenwei Zhang, Xiaohui Yan, Yi Chang and Philip Yu .....	3090
<i>Hierarchical Neural Networks for Sequential Sentence Classification in Medical Scientific Abstracts</i> Di Jin and Peter Szolovits .....	3100
<i>Investigating Capsule Networks with Dynamic Routing for Text Classification</i> Min Yang, Wei Zhao, Jianbo Ye, Zeyang Lei, Zhou Zhao and Soufei Zhang .....	3110
<i>Topic Memory Networks for Short Text Classification</i> Jichuan Zeng, Jing Li, Yan Song, Cuiyun Gao, Michael R. Lyu and Irwin King .....	3120
<i>Few-Shot and Zero-Shot Multi-Label Learning for Structured Label Spaces</i> Anthony Rios and Ramakanth Kavuluru .....	3132
<i>Automatic Poetry Generation with Mutual Reinforcement Learning</i> Xiaoyuan Yi, Maosong Sun, Ruoyu Li and Wenhao Li .....	3143

<i>Variational Autoregressive Decoder for Neural Response Generation</i>	
Jiacheng Du, Wenjie Li, Yulan He, Ruifeng Xu, Lidong Bing and Xuan Wang .....	3154
<i>Integrating Transformer and Paraphrase Rules for Sentence Simplification</i>	
Sanqiang Zhao, Rui Meng, Daqing He, Andi Saptono and Bambang Parmanto .....	3164
<i>Learning Neural Templates for Text Generation</i>	
Sam Wiseman, Stuart Shieber and Alexander Rush .....	3174
<i>Multi-Reference Training with Pseudo-References for Neural Translation and Text Generation</i>	
Renjie Zheng, Mingbo Ma and Liang Huang .....	3188
<i>Knowledge Graph Embedding with Hierarchical Relation Structure</i>	
Zhao Zhang, Fuzhen Zhuang, Meng Qu, Fen Lin and Qing He .....	3198
<i>Embedding Multimodal Relational Data for Knowledge Base Completion</i>	
Pouya Pezeshkpour, Liyan Chen and Sameer Singh .....	3208
<i>Multi-Task Identification of Entities, Relations, and Coreference for Scientific Knowledge Graph Construction</i>	
Yi Luan, Luheng He, Mari Ostendorf and Hannaneh Hajishirzi .....	3219
<i>Playing 20 Question Game with Policy-Based Reinforcement Learning</i>	
Huang Hu, Xianchao Wu, Bingfeng Luo, Chongyang Tao, Can Xu, wei wu and Zhan Chen ..	3233
<i>Multi-Hop Knowledge Graph Reasoning with Reward Shaping</i>	
Xi Victoria Lin, Richard Socher and Caiming Xiong .....	3243
<i>Neural Transductive Learning and Beyond: Morphological Generation in the Minimal-Resource Setting</i>	
Katharina Kann and Hinrich Schütze .....	3254
<i>Implicational Universals in Stochastic Constraint-Based Phonology</i>	
Giorgio Magri .....	3265
<i>Explaining Character-Aware Neural Networks for Word-Level Prediction: Do They Discover Linguistic Rules?</i>	
Frédéric Godin, Kris Demuyck, Joni Dambre, Wesley De Neve and Thomas Demeester ...	3275
<i>Adapting Word Embeddings to New Languages with Morphological and Phonological Subword Representations</i>	
Aditi Chaudhary, Chunting Zhou, Lori Levin, Graham Neubig, David R. Mortensen and Jaime Carbonell .....	3285
<i>A Computational Exploration of Exaggeration</i>	
Enrica Troiano, Carlo Strapparava, Gözde Özbal and Serra Sinem Tekiroglu .....	3296
<i>Building Context-aware Clause Representations for Situation Entity Type Classification</i>	
Zeyu Dai and Ruihong Huang .....	3305
<i>Hierarchical Dirichlet Gaussian Marked Hawkes Process for Narrative Reconstruction in Continuous Time Domain</i>	
Yeon Seonwoo, Alice Oh and Sungjoon Park .....	3316
<i>Investigating the Role of Argumentation in the Rhetorical Analysis of Scientific Publications with Neural Multi-Task Learning Models</i>	
Anne Lauscher, Goran Glavaš, Simone Paolo Ponzetto and Kai Eckert .....	3326

<i>Neural Ranking Models for Temporal Dependency Structure Parsing</i> Yuchen Zhang and Nianwen Xue .....	3339
<i>Causal Explanation Analysis on Social Media</i> Youngseo Son, Nipun Bayas and H. Andrew Schwartz .....	3350
<i>LRMM: Learning to Recommend with Missing Modalities</i> Cheng Wang, Mathias Niepert and Hui Li .....	3360
<i>Content Explorer: Recommending Novel Entities for a Document Writer</i> Michal Lukasik and Richard Zens .....	3371
<i>A Genre-Aware Attention Model to Improve the Likability Prediction of Books</i> Suraj Maharjan, Manuel Montes, Fabio A. González and Thamar Solorio .....	3381
<i>Thread Popularity Prediction and Tracking with a Permutation-invariant Model</i> Hou Pong Chan and Irwin King .....	3392
<i>IARM: Inter-Aspect Relation Modeling with Memory Networks in Aspect-Based Sentiment Analysis</i> Navonil Majumder, Soujanya Poria, Alexander Gelbukh, Md Shad Akhtar, Erik Cambria and Asif Ekbal .....	3402
<i>Limbic: Author-Based Sentiment Aspect Modeling Regularized with Word Embeddings and Discourse Relations</i> Zhe Zhang and Munindar Singh .....	3412
<i>An Interpretable Neural Network with Topical Information for Relevant Emotion Ranking</i> Yang Yang, Deyu ZHOU and Yulan He .....	3423
<i>Multi-grained Attention Network for Aspect-Level Sentiment Classification</i> Feifan Fan, Yansong Feng and Dongyan Zhao .....	3433
<i>Attentive Gated Lexicon Reader with Contrastive Contextual Co-Attention for Sentiment Classification</i> Yi Tay, Anh Tuan Luu, Siu Cheung Hui and Jian Su .....	3443
<i>Contextual Inter-modal Attention for Multi-modal Sentiment Analysis</i> Deepanway Ghosal, Md Shad Akhtar, Dushyant Chauhan, Soujanya Poria, Asif Ekbal and Pushpak Bhattacharyya .....	3454
<i>Adaptive Semi-supervised Learning for Cross-domain Sentiment Classification</i> Ruidan He, Wee Sun Lee, Hwee Tou Ng and Daniel Dahlmeier .....	3467
<i>ExtRA: Extracting Prominent Review Aspects from Customer Feedback</i> Zhiyi Luo, Shanshan Huang, Frank F. Xu, Bill Yuchen Lin, Hanyuan Shi and Kenny Zhu .....	3477
<i>Cross-Lingual Cross-Platform Rumor Verification Pivoting on Multimedia Content</i> Weiming Wen, Songwen Su and Zhou Yu .....	3487
<i>Extractive Adversarial Networks: High-Recall Explanations for Identifying Personal Attacks in Social Media Posts</i> Samuel Carton, Qiaozhu Mei and Paul Resnick .....	3497
<i>Automatic Detection of Vague Words and Sentences in Privacy Policies</i> Logan Lebanoff and Fei Liu .....	3508



<i>Multi-view Models for Political Ideology Detection of News Articles</i> Vivek Kulkarni, Junting Ye, Steve Skiena and William Yang Wang .....	3518
<i>Predicting Factuality of Reporting and Bias of News Media Sources</i> Ramy Baly, Georgi Karadzhov, Dimitar Alexandrov, James Glass and Preslav Nakov .....	3528
<i>Legal Judgment Prediction via Topological Learning</i> Haoxi Zhong, Guo Zhipeng, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu and Maosong Sun .....	3540
<i>Hierarchical CVAE for Fine-Grained Hate Speech Classification</i> Jing Qian, Mai ElSherief, Elizabeth Belding and William Yang Wang .....	3550
<i>Residualized Factor Adaptation for Community Social Media Prediction Tasks</i> Mohammadzaman Zamani, H. Andrew Schwartz, Veronica Lynn, Salvatore Giorgi and Niranjan Balasubramanian .....	3560
<i>Framing and Agenda-setting in Russian News: a Computational Analysis of Intricate Political Strategies</i> Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky and Yulia Tsvetkov ..	3570
<i>Identifying the sentiment styles of YouTube’s vloggers</i> Bennett Kleinberg, Maximilian Mozes and Isabelle van der Vegt .....	3581
<i>Native Language Identification with User Generated Content</i> Gili Goldin, Ella Rabinovich and Shuly Wintner .....	3591
<i>Beyond Error Propagation in Neural Machine Translation: Characteristics of Language Also Matter</i> Lijun Wu, Xu Tan, Di He, Fei Tian, Tao Qin, Jianhuang Lai and Tie-Yan Liu .....	3602
<i>A Study of Reinforcement Learning for Neural Machine Translation</i> Lijun Wu, Fei Tian, Tao Qin, Jianhuang Lai and Tie-Yan Liu .....	3612
<i>Meta-Learning for Low-Resource Neural Machine Translation</i> Jiatao Gu, Yong Wang, Yun Chen, Victor O. K. Li and Kyunghyun Cho .....	3622
<i>Unsupervised Statistical Machine Translation</i> Mikel Artetxe, Gorka Labaka and Eneko Agirre .....	3632
<i>A Visual Attention Grounding Neural Model for Multimodal Machine Translation</i> Mingyang Zhou, Runxiang Cheng, Yong Jae Lee and Zhou Yu .....	3643
<i>Sentiment Classification towards Question-Answering with Hierarchical Matching Network</i> Chenlin Shen, Changlong Sun, Jingjing Wang, Yangyang Kang, Shoushan Li, Xiaozhong Liu, Luo Si, Min Zhang and Guodong Zhou .....	3654
<i>Cross-topic Argument Mining from Heterogeneous Sources</i> Christian Stab, Tristan Miller, Benjamin Schiller, Pranav Rai and Iryna Gurevych .....	3664
<i>Summarizing Opinions: Aspect Extraction Meets Sentiment Prediction and They Are Both Weakly Supervised</i> Stefanos Angelidis and Mirella Lapata .....	3675
<i>CARER: Contextualized Affect Representations for Emotion Recognition</i> Elvis Saravia, Hsien-Chi Toby Liu, Yen-Hao Huang, Junlin Wu and Yi-Shin Chen .....	3687

<i>Noise Contrastive Estimation and Negative Sampling for Conditional Models: Consistency and Statistical Efficiency</i>	
Zhuang Ma and Michael Collins .....	3698
<i>CaLcs: Continuously Approximating Longest Common Subsequence for Sequence Level Optimization</i>	
Semih Yavuz, Chung-Cheng Chiu, Patrick Nguyen and Yonghui Wu .....	3708
<i>Pathologies of Neural Models Make Interpretations Difficult</i>	
Shi Feng, Eric Wallace, Alvin Grissom II, Mohit Iyyer, Pedro Rodriguez and Jordan Boyd-Graber	
3719	
<i>Phrase-level Self-Attention Networks for Universal Sentence Encoding</i>	
Wei Wu, Houfeng Wang, Tianyu Liu and Shuming Ma .....	3729
<i>BanditSum: Extractive Summarization as a Contextual Bandit</i>	
Yue Dong, Yikang Shen, Eric Crawford, Herke van Hoof and Jackie Chi Kit Cheung .....	3739
<i>A Word-Complexity Lexicon and A Neural Readability Ranking Model for Lexical Simplification</i>	
Mounica Maddela and Wei Xu .....	3749
<i>Learning Latent Semantic Annotations for Grounding Natural Language to Structured Data</i>	
Guanghui Qin, Jin-Ge Yao, Xuening Wang, Jinpeng Wang and Chin-Yew Lin .....	3761
<i>Syntactic Scaffolds for Semantic Structures</i>	
Swabha Swayamdipta, Sam Thomson, Kenton Lee, Luke Zettlemoyer, Chris Dyer and Noah A. Smith .....	3772
<i>Hierarchical Quantized Representations for Script Generation</i>	
Noah Weber, Leena Shekhar, Niranjan Balasubramanian and Nate Chambers .....	3783
<i>Semantic Role Labeling for Learner Chinese: the Importance of Syntactic Parsing and L2-L1 Parallel Data</i>	
Zi Lin, Yuguang Duan, Yuanyuan Zhao, Weiwei Sun and Xiaojun Wan .....	3793
<i>A Teacher-Student Framework for Maintainable Dialog Manager</i>	
Weikang Wang, Jiajun Zhang, Han Zhang, Mei-Yuh Hwang, Chengqing Zong and Zhifei Li .	3803
<i>Discriminative Deep Dyna-Q: Robust Planning for Dialogue Policy Learning</i>	
Shang-Yu Su, Xiujun Li, Jianfeng Gao, Jingjing Liu and Yun-Nung Chen .....	3813
<i>A Self-Attentive Model with Gate Mechanism for Spoken Language Understanding</i>	
Changliang Li, Liang Li and Ji Qi .....	3824
<i>Learning End-to-End Goal-Oriented Dialog with Multiple Answers</i>	
Janarthanan Rajendran, Jatin Ganhotra, Satinder Singh and Lazaros Polymenakos .....	3834
<i>AirDialogue: An Environment for Goal-Oriented Dialogue Research</i>	
Wei Wei, Quoc Le, Andrew Dai and Jia Li .....	3844
<i>QuaSE: Sequence Editing under Quantifiable Guidance</i>	
Yi Liao, Lidong Bing, Piji Li, Shuming Shi, Wai Lam and Tong Zhang .....	3855
<i>Paraphrase Generation with Deep Reinforcement Learning</i>	
Zichao Li, Xin Jiang, Lifeng Shang and Hang Li .....	3865

<i>Operation-guided Neural Networks for High Fidelity Data-To-Text Generation</i> Feng Nie, Jinpeng Wang, Jin-Ge Yao, Rong Pan and Chin-Yew Lin .....	3879
<i>Generating Classical Chinese Poems via Conditional Variational Autoencoder and Adversarial Training</i> Juntao Li, Yan Song, Haisong Zhang, Dongmin Chen, Shuming Shi, Dongyan Zhao and Rui Yan	3890
<i>Paragraph-level Neural Question Generation with Maxout Pointer and Gated Self-attention Networks</i> Yao Zhao, Xiaochuan Ni, Yuanyuan Ding and Qifa Ke .....	3901
<i>Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task</i> Tao Yu, Rui Zhang, Kai Yang, Michihiro Yasunaga, Dongxu Wang, Zifan Li, James Ma, Irene Li, Qingning Yao, Shanelle Roman, Zilin Zhang and Dragomir Radev .....	3911
<i>Unsupervised Natural Language Generation with Denoising Autoencoders</i> Markus Freitag and Scott Roy .....	3922
<i>Answer-focused and Position-aware Neural Question Generation</i> Xingwu Sun, Jing Liu, Yajuan Lyu, Wei He, Yanjun Ma and Shi Wang .....	3930
<i>Diversity-Promoting GAN: A Cross-Entropy Based Generative Adversarial Network for Diversified Text Generation</i> Jingjing Xu, Xuancheng Ren, Junyang Lin and Xu Sun .....	3940
<i>Towards a Better Metric for Evaluating Question Generation Systems</i> Preksha Nema and Mitesh M. Khapra .....	3950
<i>Stylistic Chinese Poetry Generation via Unsupervised Style Disentanglement</i> Cheng Yang, Maosong Sun, Xiaoyuan Yi and Wenhao Li .....	3960
<i>Generating More Interesting Responses in Neural Conversation Models with Distributional Constraints</i> Ashutosh Baheti, Alan Ritter, Jiwei Li and Bill Dolan .....	3970
<i>Better Conversations by Modeling, Filtering, and Optimizing for Coherence and Diversity</i> Xinnuo Xu, Ondřej Dušek, Ioannis Konstas and Verena Rieser .....	3981
<i>Incorporating Background Knowledge into Video Description Generation</i> Spencer Whitehead, Heng Ji, Mohit Bansal, Shih-Fu Chang and Clare Voss .....	3992
<i>Multimodal Differential Network for Visual Question Generation</i> Badri Narayana Patro, Sandeep Kumar, Vinod Kumar Kurmi and Vinay Namboodiri .....	4002
<i>Entity-aware Image Caption Generation</i> Di Lu, Spencer Whitehead, Lifu Huang, Heng Ji and Shih-Fu Chang .....	4013
<i>Learning to Describe Differences Between Pairs of Similar Images</i> Harsh Jhamtani and Taylor Berg-Kirkpatrick .....	4024
<i>Object Hallucination in Image Captioning</i> Anna Rohrbach, Lisa Anne Hendricks, Kaylee Burns, Trevor Darrell and Kate Saenko .....	4035
<i>Abstractive Text-Image Summarization Using Multi-Modal Attentional Hierarchical RNN</i> Jingqiang Chen and Hai Zhuge .....	4046

<i>Keyphrase Generation with Correlation Constraints</i>	
Jun Chen, Xiaoming Zhang, Yu Wu, Zhao Yan and Zhoujun Li .....	4057
<i>Closed-Book Training to Improve Summarization Encoder Memory</i>	
Yichen Jiang and Mohit Bansal .....	4067
<i>Improving Neural Abstractive Document Summarization with Structural Regularization</i>	
Wei Li, Xinyan Xiao, Yajuan Lyu and Yuanzhuo Wang .....	4078
<i>Iterative Document Representation Learning Towards Summarization with Polishing</i>	
Xiuying Chen, Shen Gao, Chongyang Tao, Yan Song, Dongyan Zhao and Rui Yan .....	4088
<i>Bottom-Up Abstractive Summarization</i>	
Sebastian Gehrmann, Yuntian Deng and Alexander Rush .....	4098
<i>Controlling Length in Abstractive Summarization Using a Convolutional Neural Network</i>	
Yizhu Liu, Zhiyi Luo and Kenny Zhu .....	4110
<i>APRIL: Interactively Learning to Summarise by Combining Active Preference Learning and Reinforcement Learning</i>	
Yang Gao, Christian M. Meyer and Iryna Gurevych .....	4120
<i>Adapting the Neural Encoder-Decoder Framework from Single to Multi-Document Summarization</i>	
Logan Lebanoff, Kaiqiang Song and Fei Liu .....	4131
<i>Semi-Supervised Learning for Neural Keyphrase Generation</i>	
Hai Ye and Lu Wang .....	4142
<i>MSMO: Multimodal Summarization with Multimodal Output</i>	
Junnan Zhu, Haoran Li, Tianshang Liu, Yu Zhou, Jiajun Zhang and Chengqing Zong .....	4154
<i>Frustratingly Easy Model Ensemble for Abstractive Summarization</i>	
Hayato Kobayashi .....	4165
<i>Automatic Pyramid Evaluation Exploiting EDU-based Extractive Reference Summaries</i>	
Tsutomu Hirao, Hidetaka Kamigaito and Masaaki Nagata .....	4177
<i>Learning to Encode Text as Human-Readable Summaries using Generative Adversarial Networks</i>	
Yaoshian Wang and Hung-yi Lee .....	4187
<i>Joint Multitask Learning for Community Question Answering Using Task-Specific Embeddings</i>	
Shafiq Joty, Lluís Màrquez and Preslav Nakov .....	4196
<i>What Makes Reading Comprehension Questions Easier?</i>	
Saku Sugawara, Kentaro Inui, Satoshi Sekine and Akiko Aizawa .....	4208
<i>Commonsense for Generative Multi-Hop Question Answering Tasks</i>	
Lisa Bauer, Yicheng Wang and Mohit Bansal .....	4220
<i>Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text</i>	
Haitian Sun, Bhuwan Dhingra, Manzil Zaheer, Kathryn Mazaitis, Ruslan Salakhutdinov and William Cohen .....	4231
<i>A Nil-Aware Answer Extraction Framework for Question Answering</i>	
Souvik Kundu and Hwee Tou Ng .....	4243

<i>Exploiting Deep Representations for Neural Machine Translation</i> Zi-Yi Dou, Zhaopeng Tu, Xing Wang, Shuming Shi and Tong Zhang .....	4253
<i>Why Self-Attention? A Targeted Evaluation of Neural Machine Translation Architectures</i> Gongbo Tang, Mathias Müller, Annette Rios and Rico Sennrich .....	4263
<i>Simplifying Neural Machine Translation with Addition-Subtraction Twin-Gated Recurrent Networks</i> Biao Zhang, Deyi Xiong, jinsong su, Qian Lin and Huiji Zhang .....	4273
<i>Speeding Up Neural Machine Translation Decoding by Cube Pruning</i> Wen Zhang, Liang Huang, Yang Feng, Lei Shen and Qun Liu .....	4284
<i>Revisiting Character-Based Neural Machine Translation with Capacity and Compression</i> Colin Cherry, George Foster, Ankur Bapna, Orhan Firat and Wolfgang Macherey .....	4295
<i>A Skeleton-Based Model for Promoting Coherence Among Sentences in Narrative Story Generation</i> Jingjing Xu, Xuancheng Ren, Yi Zhang, Qi Zeng, Xiaoyan Cai and Xu Sun .....	4306
<i>NEXUS Network: Connecting the Preceding and the Following in Dialogue Generation</i> Xiaoyu Shen, Hui Su, Wenjie Li and Dietrich Klakow .....	4316
<i>A Neural Local Coherence Model for Text Quality Assessment</i> Mohsen Mesgar and Michael Strube .....	4328
<i>Deep Attentive Sentence Ordering Network</i> Baiyun Cui, Yingming Li, Ming Chen and Zhongfei Zhang .....	4340
<i>Getting to "Hearer-old": Charting Referring Expressions Across Time</i> Ieva Staliūnaitė, Hannah Rohde, Bonnie Webber and Annie Louis .....	4350
<i>Making "fetch" happen: The influence of social and linguistic context on nonstandard word growth and decline</i> Ian Stewart and Jacob Eisenstein .....	4360
<i>Analyzing Correlated Evolution of Multiple Features Using Latent Representations</i> Yugo Murawaki .....	4371
<i>Capturing Regional Variation with Distributed Place Representations and Geographic Retrofitting</i> Dirk Hovy and Christoph Purschke .....	4383
<i>Characterizing Interactions and Relationships between People</i> Farzana Rashid and Eduardo Blanco .....	4395
<i>Why Swear? Analyzing and Inferring the Intentions of Vulgar Expressions</i> Eric Holgate, Isabel Cachola, Daniel Preoțiuc-Pietro and Junyi Jessy Li .....	4405
<i>Is it Time to Swish? Comparing Deep Learning Activation Functions Across NLP tasks</i> Steffen Eger, Paul Youssef and Iryna Gurevych .....	4415
<i>Hard Non-Monotonic Attention for Character-Level Transduction</i> Shijie Wu, Pamela Shapiro and Ryan Cotterell .....	4425
<i>Speed Reading: Learning to Read ForBackward via Shuttle</i> Tsu-Jui Fu and Wei-Yun Ma .....	4439

<i>Modeling Localness for Self-Attention Networks</i>	
Baosong Yang, Zhaopeng Tu, Derek F. Wong, Fandong Meng, Lidia S. Chao and Tong Zhang	4449
<i>Chargrid: Towards Understanding 2D Documents</i>	
Anoop R Katti, Christian Reisswig, Cordula Guder, Sebastian Brarda, Steffen Bickel, Johannes Höhne and Jean Baptiste Faddoul	4459
<i>Simple Recurrent Units for Highly Parallelizable Recurrence</i>	
Tao Lei, Yu Zhang, Sida I. Wang, Hui Dai and Yoav Artzi	4470
<i>NPRF: A Neural Pseudo Relevance Feedback Framework for Ad-hoc Information Retrieval</i>	
Canjia Li, Yingfei Sun, Ben He, Le Wang, Kai Hui, Andrew Yates, Le Sun and Jungang Xu	4482
<i>Co-Stack Residual Affinity Networks with Multi-level Attention Refinement for Matching Text Sequences</i>	
Yi Tay, Anh Tuan Luu and Siu Cheung Hui	4492
<i>Spherical Latent Spaces for Stable Variational Autoencoders</i>	
Jiacheng Xu and Greg Durrett	4503
<i>Learning Universal Sentence Representations with Mean-Max Attention Autoencoder</i>	
Minghua Zhang, Yunfang Wu, Weikang Li and Wei Li	4514
<i>Word Mover's Embedding: From Word2Vec to Document Embedding</i>	
Lingfei Wu, Ian En-Hsu Yen, Kun Xu, Fangli Xu, Avinash Balakrishnan, Pin-Yu Chen, Pradeep Ravikumar and Michael J. Witbrock	4524
<i>Multilingual Clustering of Streaming News</i>	
Sebastião Miranda, Arturs Znotins, Shay B. Cohen and Guntis Barzdins	4535
<i>Multi-Task Label Embedding for Text Classification</i>	
Honglun Zhang, Liqiang Xiao, Wenqing Chen, Yongkun Wang and Yaohui Jin	4545
<i>Semantic-Unit-Based Dilated Convolution for Multi-Label Text Classification</i>	
Junyong Lin, Qi Su, Pengcheng Yang, Shuming Ma and Xu Sun	4554
<i>MCapsNet: Capsule Network for Text with Multi-Task Learning</i>	
Liqiang Xiao, Honglun Zhang, Wenqing Chen, Yongkun Wang and Yaohui Jin	4565
<i>Uncertainty-aware generative models for inferring document class prevalence</i>	
Katherine Keith and Brendan O'Connor	4575
<i>Challenges of Using Text Classifiers for Causal Inference</i>	
Zach Wood-Doughty, Ilya Shpitser and Mark Dredze	4586
<i>Direct Output Connection for a High-Rank Language Model</i>	
Sho Takase, Jun Suzuki and Masaaki Nagata	4599
<i>Disfluency Detection using Auto-Correlational Neural Networks</i>	
Paria Jamshid Lou, Peter Anderson and Mark Johnson	4610
<i>Pyramidal Recurrent Unit for Language Modeling</i>	
Sachin Mehta, Rik Koncel-Kedziorski, Mohammad Rastegari and Hannaneh Hajishirzi	4620
<i>On Tree-Based Neural Sentence Modeling</i>	
Haoyue Shi, Hao Zhou, Jiaze Chen and Lei Li	4631

<i>Language Modeling with Sparse Product of Sememe Experts</i> Yihong Gu, Jun Yan, Hao Zhu, Zhiyuan Liu, Ruobing Xie, Maosong Sun, Fen Lin and Leyu Lin	4642
<i>Siamese Network-Based Supervised Topic Modeling</i> Minghui Huang, Yanghui Rao, Yuwei Liu, Haoran Xie and Fu Lee Wang	4652
<i>GraphBTM: Graph Enhanced Autoencoded Variational Inference for Biterm Topic Model</i> Qile Zhu, Zheng Feng and Xiaolin Li	4663
<i>Modeling Online Discourse with Coupled Distributed Topics</i> Akshay Srivatsan, Zachary Wojtowicz and Taylor Berg-Kirkpatrick	4673
<i>Learning Disentangled Representations of Texts with Application to Biomedical Abstracts</i> Sarthak Jain, Edward Banner, Jan-Willem van de Meent, Iain J Marshall and Byron C. Wallace	4683
<i>Multi-Source Domain Adaptation with Mixture of Experts</i> Jiang Guo, Darsh Shah and Regina Barzilay	4694
<i>A Neural Model of Adaptation in Reading</i> Marten van Schijndel and Tal Linzen	4704
<i>Understanding Deep Learning Performance through an Examination of Test Set Difficulty: A Psychometric Case Study</i> John Lalor, Hao Wu, Tsendsuren Munkhdalai and Hong Yu	4711
<i>Lexicosyntactic Inference in Neural Models</i> Aaron Steven White, Rachel Rudinger, Kyle Rawlins and Benjamin Van Durme	4717
<i>Dual Fixed-Size Ordinally Forgetting Encoding (FOFE) for Competitive Neural Language Models</i> Sedtawut Watcharawittayakul, Mingbin Xu and Hui Jiang	4725
<i>The Importance of Being Recurrent for Modeling Hierarchical Structure</i> Ke Tran, Arianna Bisazza and Christof Monz	4731
<i>Joint Learning for Targeted Sentiment Analysis</i> Dehong Ma, Sujian Li and Houfeng Wang	4737
<i>Revisiting the Importance of Encoding Logic Rules in Sentiment Classification</i> Kalpesh Krishna, Preethi Jyothi and Mohit Iyyer	4743
<i>A Co-Attention Neural Network Model for Emotion Cause Analysis with Emotional Context Awareness</i> Xiangju Li, Kaisong Song, Shi Feng, Daling Wang and Yifei Zhang	4752
<i>Modeling Empathy and Distress in Reaction to News Stories</i> Sven Buechel, Anneke Buffone, Barry Slaff, Lyle Ungar and Joao Sedoc	4758
<i>Interpretable Emoji Prediction via Label-Wise Attention LSTMs</i> Francesco Barbieri, Luis Espinosa Anke, Jose Camacho-Collados, Steven Schockaert and Horacio Saggion	4766
<i>A Tree-based Decoder for Neural Machine Translation</i> Xinyi Wang, Hieu Pham, Pengcheng Yin and Graham Neubig	4772
<i>Greedy Search with Probabilistic N-gram Matching for Neural Machine Translation</i> Chenze Shao, Xilin Chen and Yang Feng	4778

<i>Exploring Recombination for Efficient Decoding of Neural Machine Translation</i> Zhisong Zhang, Rui Wang, Masao Utiyama, Eiichiro Sumita and Hai Zhao . . . . .	4785
<i>Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation</i> Samuel Läubli, Rico Sennrich and Martin Volk . . . . .	4791
<i>Automatic Reference-Based Evaluation of Pronoun Translation Misses the Point</i> Liane Guillou and Christian Hardmeier . . . . .	4797
<i>FewRel: A Large-Scale Supervised Few-Shot Relation Classification Dataset with State-of-the-Art Evaluation</i> Xu Han, Hao Zhu, Pengfei Yu, Ziyun Wang, Yuan Yao, Zhiyuan Liu and Maosong Sun . . . . .	4803
<i>A strong baseline for question relevancy ranking</i> Ana Gonzalez, Isabelle Augenstein and Anders Søgaard . . . . .	4810
<i>Learning Sequence Encoders for Temporal Knowledge Graph Completion</i> Alberto Garcia-Duran, Sebastijan Dumančić and Mathias Niepert . . . . .	4816
<i>Similar but not the Same: Word Sense Disambiguation Improves Event Detection via Neural Representation Matching</i> Weiyi Lu and Thien Huu Nguyen . . . . .	4822
<i>Learning Word Representations with Cross-Sentence Dependency for End-to-End Co-reference Resolution</i> Hongyin Luo and Jim Glass . . . . .	4829
<i>Word Relation Autoencoder for Unseen Hypernym Extraction Using Word Embeddings</i> Hong-You Chen, Cheng-Syuan Lee, Keng-Te Liao and Shou-de Lin . . . . .	4834
<i>Refining Pretrained Word Embeddings Using Layer-wise Relevance Propagation</i> Akira Utsumi . . . . .	4840
<i>Learning Gender-Neutral Word Embeddings</i> Jieyu Zhao, Yichao Zhou, Zeyu Li, Wei Wang and Kai-Wei Chang . . . . .	4847
<i>Learning Concept Abstractness Using Weak Supervision</i> Ella Rabinovich, Benjamin Sznajder, Artem Spector, Ilya Shnayderman, Ranit Aharonov, David Konopnicki and Noam Slonim . . . . .	4854
<i>Word Sense Induction with Neural biLM and Symmetric Patterns</i> Asaf Amrami and Yoav Goldberg . . . . .	4860
<i>InferLite: Simple Universal Sentence Representations from Natural Language Inference Data</i> Jamie Kiros and William Chan . . . . .	4868
<i>Similarity-Based Reconstruction Loss for Meaning Representation</i> Olga Kovaleva, Anna Rumshisky and Alexey Romanov . . . . .	4875
<i>What can we learn from Semantic Tagging?</i> Mostafa Abdou, Artur Kulmizev, Vinit Ravishankar, Lasha Abzianidze and Johan Bos . . . . .	4881
<i>Conditional Word Embedding and Hypothesis Testing via Bayes-by-Backprop</i> Rujun Han, Michael Gill, Arthur Spirling and Kyunghyun Cho . . . . .	4890



<i>Classifying Referential and Non-referential It Using Gaze</i> Victoria Yaneva, Le An Ha, Richard Evans and Ruslan Mitkov .....	4896
<i>State-of-the-art Chinese Word Segmentation with Bi-LSTMs</i> Ji Ma, Kuzman Ganchev and David Weiss .....	4902
<i>Sanskrit Sandhi Splitting using seq2(seq)2</i> Rahul Aralikatte, Neelamadhav Gantayat, Naveen Panwar, Anush Sankaran and Senthil Mani	4909
<i>Unsupervised Neural Word Segmentation for Chinese via Segmental Language Modeling</i> Zhiqing Sun and Zhi-Hong Deng .....	4915
<i>LemmaTag: Jointly Tagging and Lemmatizing for Morphologically Rich Languages with BRNNs</i> Daniel Kondratyuk, Tomáš Gavenčiak, Milan Straka and Jan Hajič .....	4921
<i>Recovering Missing Characters in Old Hawaiian Writing</i> Brendan Shillingford and Oiwi Parker Jones .....	4929
<i>When data permutations are pathological: the case of neural natural language inference</i> Natalie Schluter and Daniel Varab .....	4935
<i>Bridging Knowledge Gaps in Neural Entailment via Symbolic Models</i> Dongyeop Kang, Tushar Khot, Ashish Sabharwal and Peter Clark .....	4940
<i>The BQ Corpus: A Large-scale Domain-specific Chinese Corpus For Sentence Semantic Equivalence Identification</i> Jing Chen, Qingcai Chen, Xin Liu, Haijun Yang, Daohe Lu and Buzhou Tang .....	4946
<i>Interpreting Recurrent and Attention-Based Neural Models: a Case Study on Natural Language Inference</i> Reza Ghaeini, Xiaoli Fern and Prasad Tadepalli .....	4952
<i>Towards Semi-Supervised Learning for Deep Semantic Role Labeling</i> Sanket Vaibhav Mehta, Jay Yoon Lee and Jaime Carbonell .....	4958
<i>Identifying Domain Adjacent Instances for Semantic Parsers</i> James Ferguson, Janara Christensen, Edward Li and Edgar González .....	4964
<i>Mapping natural language commands to web elements</i> Panupong Pasupat, Tian-Shun Jiang, Evan Liu, Kelvin Guu and Percy Liang .....	4970
<i>Wronging a Right: Generating Better Errors to Improve Grammatical Error Detection</i> Sudhanshu Kasewa, Pontus Stenetorp and Sebastian Riedel .....	4977
<i>Modeling Input Uncertainty in Neural Network Dependency Parsing</i> Rob van der Goot and Gertjan van Noord .....	4984
<i>Parameter sharing between dependency parsers for related languages</i> Miryam de Lhoneux, Johannes Bjerva, Isabelle Augenstein and Anders Søgaard .....	4992
<i>Grammar Induction with Neural Language Models: An Unusual Replication</i> Phu Mon Htut, Kyunghyun Cho and Samuel Bowman .....	4998
<i>Data Augmentation via Dependency Tree Morphing for Low-Resource Languages</i> Gozde Gul Sahin and Mark Steedman .....	5004

<i>How Much Reading Does Reading Comprehension Require? A Critical Investigation of Popular Benchmarks</i>	
Divyansh Kaushik and Zachary C. Lipton . . . . .	5010
<i>MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling</i>	
Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan and Milica Gasic . . . . .	5016
<i>Linguistically-Informed Self-Attention for Semantic Role Labeling</i>	
Emma Strubell, Patrick Verga, Daniel Andor, David Weiss and Andrew McCallum . . . . .	5027
<i>Phrase-Based &amp; Neural Unsupervised Machine Translation</i>	
Guillaume Lample, Myle Ott, Alexis Conneau, Ludovic Denoyer and Marc’Aurelio Ranzato .	5039