

Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)

Online
16 – 20 November 2020

Volume 1 of 15

ISBN: 978-1-7138-1979-0

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© (2020) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

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>Detecting Attackable Sentences in Arguments</i>	
Yohan Jo, Seojin Bang, Emaad Manzoor, Eduard Hovy and Chris Reed	1
<i>Extracting Implicitly Asserted Propositions in Argumentation</i>	
Yohan Jo, Jacky Visser, Chris Reed and Eduard Hovy	24
<i>Quantitative Argument Summarization and Beyond: Cross-Domain Key Point Analysis</i>	
Roy Bar-Haim, Yoav Kantor, Lilach Eden, Roni Friedman, Dan Lahav and Noam Slonim	39
<i>Unsupervised Stance Detection for Arguments from Consequences</i>	
Jonathan Kobbe, Ioana Hulpuş and Heiner Stuckenschmidt	50
<i>BLEU might be Guilty but References are not Innocent</i>	
Markus Freitag, David Grangier and Isaac Caswell	61
<i>Translationese in Machine Translation Evaluation</i>	
Yvette Graham, Barry Haddow and Philipp Koehn	72
<i>Simulated Multiple Reference Training Improves Low-Resource Machine Translation</i>	
Huda Khayrallah, Brian Thompson, Matt Post and Philipp Koehn	82
<i>Automatic Machine Translation Evaluation in Many Languages via Zero-Shot Paraphrasing</i>	
Brian Thompson and Matt Post	90
<i>PRover: Proof Generation for Interpretable Reasoning over Rules</i>	
Swarnadeep Saha, Sayan Ghosh, Shashank Srivastava and Mohit Bansal	122
<i>Learning to Explain: Datasets and Models for Identifying Valid Reasoning Chains in Multihop Question-Answering</i>	
Harsh Jhamtani and Peter Clark	137
<i>Self-Supervised Knowledge Triplet Learning for Zero-shot Question Answering</i>	
Pratyay Banerjee and Chitta Baral	151
<i>More Bang for Your Buck: Natural Perturbation for Robust Question Answering</i>	
Daniel Khashabi, Tushar Khot and Ashish Sabharwal	163
<i>A Matter of Framing: The Impact of Linguistic Formalism on Probing Results</i>	
Iliia Kuznetsov and Iryna Gurevych	171
<i>Information-Theoretic Probing with Minimum Description Length</i>	
Elena Voita and Ivan Titov	183
<i>Intrinsic Probing through Dimension Selection</i>	
Lucas Torroba Hennigen, Adina Williams and Ryan Cotterell	197
<i>Learning Helpful Inductive Biases from Self-Supervised Pretraining</i>	
Alex Warstadt, Yian Zhang, Xiaocheng Li, Haokun Liu and Samuel R. Bowman	217
<i>Repulsive Attention: Rethinking Multi-head Attention as Bayesian Inference</i>	
Bang An, Jie Lyu, Zhenyi Wang, Chunyuan Li, Changwei Hu, Fei Tan, Ruiyi Zhang, Yifan Hu and Changyou Chen	236

<i>KERMIT: Complementing Transformer Architectures with Encoders of Explicit Syntactic Interpretations</i> Fabio Massimo Zanzotto, Andrea Santilli, Leonardo Ranaldi, Dario Onorati, Pierfrancesco Tommasino and Francesca Fallucchi	256
<i>ETC: Encoding Long and Structured Inputs in Transformers</i> Joshua Ainslie, Santiago Ontanon, Chris Alberti, Vaclav Cvicek, Zachary Fisher, Philip Pham, Anirudh Ravula, Sumit Sanghai, Qifan Wang and Li Yang	268
<i>Pre-Training Transformers as Energy-Based Cloze Models</i> Kevin Clark, Minh-Thang Luong, Quoc Le and Christopher D. Manning	285
<i>Calibration of Pre-trained Transformers</i> Shrey Desai and Greg Durrett	295
<i>Near-imperceptible Neural Linguistic Steganography via Self-Adjusting Arithmetic Coding</i> Jiaming Shen, Heng Ji and Jiawei Han	303
<i>Multi-Dimensional Gender Bias Classification</i> Emily Dinan, Angela Fan, Ledell Wu, Jason Weston, Douwe Kiela and Adina Williams	314
<i>Human-in-the-loop Debugging Deep Text Classifiers</i> Piyawat Lertvittayakumjorn, Lucia Specia and Francesca Toni	332
<i>Conversational Document Prediction to Assist Customer Care Agents</i> Jatin Ganhotra, Haggai Roitman, Doron Cohen, Nathaniel Mills, Chulaka Gunasekara, Yosi Mass, Sachindra Joshi, Luis Lastras and David Konopnicki	349
<i>Incremental Processing in the Age of Non-Incremental Encoders: An Empirical Assessment of Bidirectional Models for Incremental NLU</i> Brielen Madureira and David Schlangen	357
<i>Augmented Natural Language for Generative Sequence Labeling</i> Ben Athiwaratkun, Cicero Nogueira dos Santos, Jason Krone and Bing Xiang	375
<i>Dialogue Response Ranking Training with Large-Scale Human Feedback Data</i> Xiang Gao, Yizhe Zhang, Michel Galley, Chris Brockett and Bill Dolan	386
<i>Semantic Evaluation for Text-to-SQL with Distilled Test Suite</i> Ruiqi Zhong, Tao Yu and Dan Klein	396
<i>Cross-Thought for Sentence Encoder Pre-training</i> Shuohang Wang, Yuwei Fang, Siqi Sun, Zhe Gan, Yu Cheng, Jingjing Liu and Jing Jiang	412
<i>AutoQA: From Databases To Q&A Semantic Parsers With Only Synthetic Training Data</i> Silei Xu, Sina Semnani, Giovanni Campagna and Monica Lam	422
<i>A Spectral Method for Unsupervised Multi-Document Summarization</i> Kexiang Wang, Baobao Chang and Zhifang Sui	435
<i>What Have We Achieved on Text Summarization?</i> Dandan Huang, Leyang Cui, Sen Yang, Guangsheng Bao, Kun Wang, Jun Xie and Yue Zhang	446
<i>Q-learning with Language Model for Edit-based Unsupervised Summarization</i> Ryosuke Kohita, Akifumi Wachi, Yang Zhao and Ryuki Tachibana	470

<i>Friendly Topic Assistant for Transformer Based Abstractive Summarization</i>	
Zhengjue Wang, Zhibin Duan, Hao Zhang, chaojie wang, long tian, Bo Chen and Mingyuan Zhou	
485	
<i>Contrastive Distillation on Intermediate Representations for Language Model Compression</i>	
Siqi Sun, Zhe Gan, Yuwei Fang, Yu Cheng, Shuohang Wang and Jingjing Liu	498
<i>TernaryBERT: Distillation-aware Ultra-low Bit BERT</i>	
Wei Zhang, Lu Hou, Yichun Yin, Lifeng Shang, Xiao Chen, Xin Jiang and Qun Liu	509
<i>Self-Supervised Meta-Learning for Few-Shot Natural Language Classification Tasks</i>	
Trapit Bansal, Rishikesh Jha, Tsendsuren Munkhdalai and Andrew McCallum	522
<i>Efficient Meta Lifelong-Learning with Limited Memory</i>	
Zirui Wang, Sanket Vaibhav Mehta, Barnabas Poczos and Jaime Carbonell	535
<i>On the Evaluation of Contextual Embeddings for Zero-Shot Cross-Lingual Transfer Learning</i>	
Phillip Keung, Yichao Lu, Julian Salazar and Vikas Bhardwaj	549
<i>A Supervised Word Alignment Method based on Cross-Language Span Prediction using Multilingual BERT</i>	
Masaaki Nagata, Katsuki Chousa and Masaaki Nishino	555
<i>Accurate Word Alignment Induction from Neural Machine Translation</i>	
Yun Chen, Yang Liu, Guanhua Chen, Xin Jiang and Qun Liu	566
<i>ChrEn: Cherokee-English Machine Translation for Endangered Language Revitalization</i>	
Shiyue Zhang, Benjamin Frey and Mohit Bansal	577
<i>Unsupervised Discovery of Implicit Gender Bias</i>	
Anjalie Field and Yulia Tsvetkov	596
<i>Condolences and Empathy in Online Communities</i>	
Naitian Zhou and David Jurgens	609
<i>An Embedding Model for Estimating Legislative Preferences from the Frequency and Sentiment of Tweets</i>	
Gregory Spell, Brian Guay, Sunshine Hillygus and Lawrence Carin	627
<i>Measuring Information Propagation in Literary Social Networks</i>	
Matthew Sims and David Bamman	642
<i>Social Chemistry 101: Learning to Reason about Social and Moral Norms</i>	
Maxwell Forbes, Jena D. Hwang, Vered Shwartz, Maarten Sap and Yejin Choi	653
<i>Event Extraction by Answering (Almost) Natural Questions</i>	
Xinya Du and Claire Cardie	671
<i>Connecting the Dots: Event Graph Schema Induction with Path Language Modeling</i>	
Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers and Clare Voss	684
<i>Joint Constrained Learning for Event-Event Relation Extraction</i>	
Haoyu Wang, Muhao Chen, Hongming Zhang and Dan Roth	696
<i>Incremental Event Detection via Knowledge Consolidation Networks</i>	
Pengfei Cao, Yubo Chen, Jun Zhao and Taifeng Wang	707

<i>Semi-supervised New Event Type Induction and Event Detection</i> Lifu Huang and Heng Ji	718
<i>Language Generation with Multi-hop Reasoning on Commonsense Knowledge Graph</i> Haozhe Ji, Pei Ke, Shaohan Huang, Furu Wei, Xiaoyan Zhu and Minlie Huang	725
<i>Reformulating Unsupervised Style Transfer as Paraphrase Generation</i> Kalpesh Krishna, John Wieting and Mohit Iyyer	737
<i>De-biased Court’s View Generation with Causality</i> Yiquan Wu, Kun Kuang, Yating Zhang, Xiaozhong Liu, Changlong Sun, Jun Xiao, Yueting Zhuang, Luo Si and Fei Wu	763
<i>PAIR: Planning and Iterative Refinement in Pre-trained Transformers for Long Text Generation</i> Xinyu Hua and Lu Wang	781
<i>Backpropagation-based Decoding for Unsupervised Counterfactual and Abductive Reasoning</i> Lianhui Qin, Vered Shwartz, Peter West, Chandra Bhagavatula, Jena D. Hwang, Ronan Le Bras, Antoine Bosselut and Yejin Choi	794
<i>Where Are You? Localization from Embodied Dialog</i> Meera Hahn, Jacob Krantz, Dhruv Batra, Devi Parikh, James Rehg, Stefan Lee and Peter Anderson	806
<i>Learning to Represent Image and Text with Denotation Graphs</i> Bowen Zhang, Hexiang Hu, Vihan Jain, Eugene Ie and Fei Sha	823
<i>Video2Commonsense: Generating Commonsense Descriptions to Enrich Video Captioning</i> Zhiyuan Fang, Tejas Gokhale, Pratyay Banerjee, Chitta Baral and Yezhou Yang	840
<i>Does my multimodal model learn cross-modal interactions? It’s harder to tell than you might think!</i> Jack Hessel and Lillian Lee	861
<i>MUTANT: A Training Paradigm for Out-of-Distribution Generalization in Visual Question Answering</i> Tejas Gokhale, Pratyay Banerjee, Chitta Baral and Yezhou Yang	878
<i>Mitigating Gender Bias for Neural Dialogue Generation with Adversarial Learning</i> Haochen Liu, Wentao Wang, Yiqi Wang, Hui Liu, Zitao Liu and Jiliang Tang	893
<i>Will I Sound Like Me? Improving Persona Consistency in Dialogues through Pragmatic Self-Consciousness</i> Hyunwoo Kim, Byeongchang Kim and Gunhee Kim	904
<i>TOD-BERT: Pre-trained Natural Language Understanding for Task-Oriented Dialogue</i> Chien-Sheng Wu, Steven C.H. Hoi, Richard Socher and Caiming Xiong	917
<i>RiSAWOZ: A Large-Scale Multi-Domain Wizard-of-Oz Dataset with Rich Semantic Annotations for Task-Oriented Dialogue Modeling</i> Jun Quan, Shian Zhang, Qian Cao, Zizhong Li and Deyi Xiong	930
<i>Filtering Noisy Dialogue Corpora by Connectivity and Content Relatedness</i> Reina Akama, Sho Yokoi, Jun Suzuki and Kentaro Inui	941
<i>Latent Geographical Factors for Analyzing the Evolution of Dialects in Contact</i> Yugo Murawaki	959

<i>Predicting Reference: What do Language Models Learn about Discourse Models?</i>	
Shiva Upadhye, Leon Bergen and Andrew Kehler	977
<i>Word class flexibility: A deep contextualized approach</i>	
Bai Li, Guillaume Thomas, Yang Xu and Frank Rudzicz	983
<i>Shallow-to-Deep Training for Neural Machine Translation</i>	
Bei Li, Ziyang Wang, Hui Liu, Yufan Jiang, Quan Du, Tong Xiao, Huizhen Wang and Jingbo Zhu	995
<i>Iterative Refinement in the Continuous Space for Non-Autoregressive Neural Machine Translation</i>	
Jason Lee, Raphael Shu and Kyunghyun Cho	1006
<i>Why Skip If You Can Combine: A Simple Knowledge Distillation Technique for Intermediate Layers</i>	
Yimeng Wu, Peyman Passban, Mehdi Rezagholizadeh and Qun Liu	1016
<i>Multi-task Learning for Multilingual Neural Machine Translation</i>	
Yiren Wang, ChengXiang Zhai and Hany Hassan	1022
<i>Token-level Adaptive Training for Neural Machine Translation</i>	
Shuhao Gu, Jinchao Zhang, Fandong Meng, Yang Feng, Wanying Xie, Jie Zhou and Dong Yu	1035
<i>Multi-Unit Transformers for Neural Machine Translation</i>	
Jianhao Yan, Fandong Meng and Jie Zhou	1047
<i>On the Sparsity of Neural Machine Translation Models</i>	
Yong Wang, Longyue Wang, Victor Li and Zhaopeng Tu	1060
<i>Incorporating a Local Translation Mechanism into Non-autoregressive Translation</i>	
Xiang Kong, Zhisong Zhang and Eduard Hovy	1067
<i>Self-Paced Learning for Neural Machine Translation</i>	
Yu Wan, Baosong Yang, Derek F. Wong, Yikai Zhou, Lidia S. Chao, Haibo Zhang and Boxing Chen	1074
<i>Long-Short Term Masking Transformer: A Simple but Effective Baseline for Document-level Neural Machine Translation</i>	
Pei Zhang, Boxing Chen, Niyu Ge and Kai Fan	1081
<i>Generating Diverse Translation from Model Distribution with Dropout</i>	
Xuanfu Wu, Yang Feng and Chenze Shao	1088
<i>Non-Autoregressive Machine Translation with Latent Alignments</i>	
Chitwan Saharia, William Chan, Saurabh Saxena and Mohammad Norouzi	1098
<i>Look at the First Sentence: Position Bias in Question Answering</i>	
Miyounk Ko, Jinhyuk Lee, Hyunjae Kim, Gangwoo Kim and Jaewoo Kang	1109
<i>ProtoQA: A Question Answering Dataset for Prototypical Common-Sense Reasoning</i>	
Michael Boratko, Xiang Li, Tim O’Gorman, Rajarshi Das, Dan Le and Andrew McCallum ..	1122
<i>IIRC: A Dataset of Incomplete Information Reading Comprehension Questions</i>	
James Ferguson, Matt Gardner, Hannaneh Hajishirzi, Tushar Khot and Pradeep Dasigi	1137
<i>Unsupervised Adaptation of Question Answering Systems via Generative Self-training</i>	
Steven Rennie, Etienne Marcheret, Neil Mallinar, David Nahamoo and Vaibhava Goel	1148

<i>TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions</i>	
Qiang Ning, Hao Wu, Rujun Han, Nanyun Peng, Matt Gardner and Dan Roth	1158
<i>ToTTo: A Controlled Table-To-Text Generation Dataset</i>	
Ankur Parikh, Xuezhi Wang, Sebastian Gehrmann, Manaal Faruqui, Bhuwan Dhingra, Diyi Yang and Dipanjan Das	1173
<i>Knowledge Graph Empowered Entity Description Generation</i>	
Liyong Cheng, Dekun Wu, Lidong Bing, Yan Zhang, Zhanming Jie, Wei Lu and Luo Si	1187
<i>Small but Mighty: New Benchmarks for Split and Rephrase</i>	
Li Zhang, Huaiyu Zhu, Siddhartha Brahma and Yunyao Li	1198
<i>Online Back-Parsing for AMR-to-Text Generation</i>	
Xuefeng Bai, Linfeng Song and Yue Zhang	1206
<i>Reading Between the Lines: Exploring Infilling in Visual Narratives</i>	
Khyathi Raghavi Chandu, Ruo-Ping Dong and Alan W Black	1220
<i>Acrostic Poem Generation</i>	
Rajat Agarwal and Katharina Kann	1230
<i>Local Additivity Based Data Augmentation for Semi-supervised NER</i>	
Jiaao Chen, Zhenghui Wang, Ran Tian, Zichao Yang and Diyi Yang	1241
<i>Grounded Compositional Outputs for Adaptive Language Modeling</i>	
Nikolaos Pappas, Phoebe Mulcaire and Noah A. Smith	1252
<i>SSMBA: Self-Supervised Manifold Based Data Augmentation for Improving Out-of-Domain Robustness</i>	
Nathan Ng, Kyunghyun Cho and Marzyeh Ghassemi	1268
<i>SetConv: A New Approach for Learning from Imbalanced Data</i>	
Yang Gao, Yi-Fan Li, Yu Lin, Charu Aggarwal and Latifur Khan	1284
<i>Scalable Multi-Hop Relational Reasoning for Knowledge-Aware Question Answering</i>	
Yanlin Feng, Xinyue Chen, Bill Yuchen Lin, Peifeng Wang, Jun Yan and Xiang Ren	1295
<i>Improving Bilingual Lexicon Induction for Low Frequency Words</i>	
Jiayi Huang, Xingyu Cai and Kenneth Church	1310
<i>Learning VAE-LDA Models with Rounded Reparameterization Trick</i>	
Runzhi Tian, Yongyi Mao and Richong Zhang	1315
<i>Calibrated Fine-Tuning for Pre-trained Language Models via Manifold Smoothing</i>	
Lingkai Kong, Haoming Jiang, Yuchen Zhuang, Jie Lyu, Tuo Zhao and Chao Zhang	1326
<i>Scaling Hidden Markov Language Models</i>	
Justin Chiu and Alexander Rush	1341
<i>Coding Textual Inputs Boosts the Accuracy of Neural Networks</i>	
Abdul Rafae Khan, Jia Xu and Weiwei Sun	1350
<i>Learning from Task Descriptions</i>	
Orion Weller, Nicholas Lourie, Matt Gardner and Matthew Peters	1361

<i>Hashtags, Emotions, and Comments: A Large-Scale Dataset to Understand Fine-Grained Social Emotions to Online Topics</i>	
Keyang Ding, Jing Li and Yuji Zhang	1376
<i>Named Entity Recognition for Social Media Texts with Semantic Augmentation</i>	
Yuyang Nie, Yuanhe Tian, Xiang Wan, Yan Song and Bo Dai	1383
<i>Predicting Stance and Rumor Veracity via Dual Hierarchical Transformer with Pretrained Encoders</i>	
Jianfei Yu, Jing Jiang, Ling Min Serena Khoo, Hai Leong Chieu and Rui Xia	1392
<i>Social Media Attributions in the Context of Water Crisis</i>	
Rupak Sarkar, Sayantan Mahinder, Hirak Sarkar and Ashiqur KhudaBukhsh	1402
<i>On the Reliability and Validity of Detecting Approval of Political Actors in Tweets</i>	
Indira Sen, Fabian Flöck and Claudia Wagner	1413
<i>Towards Medical Machine Reading Comprehension with Structural Knowledge and Plain Text</i>	
Dongfang Li, Baotian Hu, Qingcai Chen, Weihua Peng and Anqi Wang	1427
<i>Generating Radiology Reports via Memory-driven Transformer</i>	
Zhihong Chen, Yan Song, Tsung-Hui Chang and Xiang Wan	1439
<i>Planning and Generating Natural and Diverse Disfluent Texts as Augmentation for Disfluency Detection</i>	
Jingfeng Yang, Diyi Yang and Zhaoran Ma	1450
<i>Predicting Clinical Trial Results by Implicit Evidence Integration</i>	
Qiao Jin, Chuanqi Tan, Moshu Chen, Xiaozhong Liu and Songfang Huang	1461
<i>Explainable Clinical Decision Support from Text</i>	
Jinyue Feng, Chantal Shaib and Frank Rudzicz	1478
<i>A Knowledge-driven Generative Model for Multi-implication Chinese Medical Procedure Entity Normalization</i>	
Jinghui Yan, Yining Wang, Lu Xiang, Yu Zhou and Chengqing Zong	1490
<i>CheXbert: Combining Automatic Labelers and Expert Annotations for Accurate Radiology Report Labeling Using BERT</i>	
Akshay Smit, Saahil Jain, Pranav Rajpurkar, Anuj Pareek, Andrew Ng and Matthew Lungren	1500
<i>Benchmarking Meaning Representations in Neural Semantic Parsing</i>	
Jiaqi Guo, Qian Liu, Jian-Guang LOU, Zhenwen Li, Xueqing Liu, Tao Xie and Ting Liu	1520
<i>Analogous Process Structure Induction for Sub-event Sequence Prediction</i>	
Hongming Zhang, Muhao Chen, Haoyu Wang, Yangqiu Song and Dan Roth	1541
<i>SLM: Learning a Discourse Language Representation with Sentence Unshuffling</i>	
Haejun Lee, Drew A. Hudson, Kangwook Lee and Christopher D. Manning	1551
<i>Detecting Fine-Grained Cross-Lingual Semantic Divergences without Supervision by Learning To Rank</i>	
Eleftheria Briakou and Marine Carpuat	1563
<i>A Bilingual Generative Transformer for Semantic Sentence Embedding</i>	
John Wieting, Graham Neubig and Taylor Berg-Kirkpatrick	1581
<i>Semantically Inspired AMR Alignment for the Portuguese language</i>	
Rafael Anchiêta and Thiago Pardo	1595

<i>An Unsupervised Sentence Embedding Method by Mutual Information Maximization</i>	
Yan Zhang, Ruidan He, ZUOZHU LIU, Kwan Hui Lim and Lidong Bing	1601
<i>Compositional Phrase Alignment and Beyond</i>	
Yuki Arase and Jun'ichi Tsujii	1611
<i>Table Fact Verification with Structure-Aware Transformer</i>	
Hongzhi Zhang, Yingyao Wang, Sirui Wang, Xuezhi Cao, Fuzheng Zhang and Zhongyuan Wang	
1624	
<i>Double Graph Based Reasoning for Document-level Relation Extraction</i>	
Shuang Zeng, Runxin Xu, Baobao Chang and Lei Li	1630
<i>Event Extraction as Machine Reading Comprehension</i>	
Jian Liu, Yubo Chen, Kang Liu, Wei Bi and Xiaojiang Liu	1641
<i>MAVEN: A Massive General Domain Event Detection Dataset</i>	
Xiaozhi Wang, Ziqi Wang, Xu Han, Wangyi Jiang, Rong Han, Zhiyuan Liu, Juanzi Li, Peng Li, Yankai Lin and Jie Zhou	1652
<i>Knowledge Graph Alignment with Entity-Pair Embedding</i>	
Zhichun Wang, Jinjian Yang and Xiaoju Ye	1672
<i>Adaptive Attentional Network for Few-Shot Knowledge Graph Completion</i>	
Jiawei Sheng, Shu Guo, Zhenyu Chen, Juwei Yue, Lihong Wang, Tingwen Liu and Hongbo Xu	
1681	
<i>Pre-training Entity Relation Encoder with Intra-span and Inter-span Information</i>	
Yijun Wang, Changzhi Sun, Yuanbin Wu, Junchi Yan, Peng Gao and Guotong Xie	1692
<i>Two are Better Than One: Joint Entity and Relation Extraction with Table-Sequence Encoders</i>	
Jue WANG and Wei Lu	1706
<i>Beyond [CLS] through Ranking by Generation</i>	
Cicero Nogueira dos Santos, Xiaofei Ma, Ramesh Nallapati, zhiheng huang and Bing Xiang .	1722
<i>Tired of Topic Models? Clusters of Pretrained Word Embeddings Make for Fast and Good Topics too!</i>	
Suzanna Sia, Ayush Dalmaia and Sabrina J. Mielke	1728
<i>Multi-document Summarization with Maximal Marginal Relevance-guided Reinforcement Learning</i>	
Yuning Mao, Yanru Qu, Yiqing Xie, Xiang Ren and Jiawei Han	1737
<i>Improving Neural Topic Models using Knowledge Distillation</i>	
Alexander Miserlis Hoyle, Pranav Goel and Philip Resnik	1752
<i>Short Text Topic Modeling with Topic Distribution Quantization and Negative Sampling Decoder</i>	
Xiaobao Wu, Chunping Li, Yan Zhu and Yishu Miao	1772
<i>Querying across genres to retrieve research that supports medical claims made in news</i>	
Chaoyuan Zuo, Narayan Acharya and Ritwik Banerjee	1783
<i>Incorporating Multimodal Information in Open-Domain Web Keyphrase Extraction</i>	
Yansen Wang, Zhen Fan and Carolyn Rose	1790

<i>MOSEAS: A Multimodal Language Dataset for Spanish, Portuguese, German and French</i> AmirAli Bagher Zadeh, Yansheng Cao, Simon Hessner, Paul Pu Liang, Soujanya Poria and Louis-Philippe Morency	1801
<i>Combining Self-Training and Self-Supervised Learning for Unsupervised Disfluency Detection</i> Shaolei Wang, Zhongyuan Wang, Wanxiang Che and Ting Liu	1813
<i>Multimodal Routing: Improving Local and Global Interpretability of Multimodal Language Analysis</i> Yao-Hung Hubert Tsai, Martin Ma, Muqiao Yang, Ruslan Salakhutdinov and Louis-Philippe Morency	1823
<i>Multistage Fusion with Forget Gate for Multimodal Summarization in Open-Domain Videos</i> Nayu Liu, Xian Sun, Hongfeng Yu, Wenkai Zhang and Guangluan Xu	1834
<i>BiST: Bi-directional Spatio-Temporal Reasoning for Video-Grounded Dialogues</i> Hung Le, Doyen Sahoo, Nancy Chen and Steven C.H. Hoi	1846
<i>UniConv: A Unified Conversational Neural Architecture for Multi-domain Task-oriented Dialogues</i> Hung Le, Doyen Sahoo, Chenghao Liu, Nancy Chen and Steven C.H. Hoi	1860
<i>GraphDialog: Integrating Graph Knowledge into End-to-End Task-Oriented Dialogue Systems</i> Shiquan Yang, Rui Zhang and Sarah Erfani	1878
<i>Structured Attention for Unsupervised Dialogue Structure Induction</i> Liang Qiu, Yizhou Zhao, Weiyan Shi, Yuan Liang, Feng Shi, Tao Yuan, Zhou Yu and Song-Chun Zhu	1889
<i>Cross Copy Network for Dialogue Generation</i> Changzhen Ji, Xin Zhou, Yating Zhang, Xiaozhong Liu, Changlong Sun, Conghui Zhu and Tiejun Zhao	1900
<i>Multi-turn Response Selection using Dialogue Dependency Relations</i> Qi Jia, Yizhu Liu, Siyu Ren, Kenny Zhu and Haifeng Tang	1911
<i>Parallel Interactive Networks for Multi-Domain Dialogue State Generation</i> Junfan Chen, Richong Zhang, Yongyi Mao and Jie Xu	1921
<i>SlotRefine: A Fast Non-Autoregressive Model for Joint Intent Detection and Slot Filling</i> Di Wu, Liang Ding, Fan Lu and Jian Xie	1932
<i>An Information Bottleneck Approach for Controlling Conciseness in Rationale Extraction</i> Bhargavi Paranjape, Mandar Joshi, John Thickstun, Hannaneh Hajishirzi and Luke Zettlemoyer	1938
<i>CrowS-Pairs: A Challenge Dataset for Measuring Social Biases in Masked Language Models</i> Nikita Nangia, Clara Vania, Rasika Bhalerao and Samuel R. Bowman	1953
<i>LOGAN: Local Group Bias Detection by Clustering</i> Jieyu Zhao and Kai-Wei Chang	1968
<i>RNNs can generate bounded hierarchical languages with optimal memory</i> John Hewitt, Michael Hahn, Surya Ganguli, Percy Liang and Christopher D. Manning	1978
<i>Detecting Independent Pronoun Bias with Partially-Synthetic Data Generation</i> Robert Munro and Alex (Carmen) Morrison	2011

<i>Visually Grounded Continual Learning of Compositional Phrases</i>	
Xisen Jin, Junyi Du, Arka Sadhu, Ram Nevatia and Xiang Ren	2018
<i>An Effective Framework for Weakly-Supervised Phrase Grounding</i>	
Qinxin Wang, Hao Tan, Sheng Shen, Michael Mahoney and Zhewei Yao	2030
<i>Finding Domain-Specific Grounding in Noisy Visual-Textual Documents</i>	
Gregory Yauney, Jack Hessel and David Mimno	2039
<i>Hero: Hierarchical Encoder for Video+Language Omni-representation Pre-training</i>	
Linjie Li, Yen-Chun Chen, Yu Cheng, Zhe Gan, Licheng Yu and Jingjing Liu	2046
<i>Vokenization: Improving Language Understanding via Contextualized, Visually-Grounded Supervision</i>	
Hao Tan and Mohit Bansal	2066
<i>Detecting Cross-Modal Inconsistency to Defend Against Neural Fake News</i>	
Reuben Tan, Bryan Plummer and Kate Saenko	2081
<i>Enhancing Aspect Term Extraction with Soft Prototypes</i>	
Zhuang Chen and Tiejun Qian	2107
<i>FedED: Federated Learning via Ensemble Distillation for Medical Relation Extraction</i>	
Dianbo Sui, Yubo Chen, Jun Zhao, Yantao Jia, Yuantao Xie and Weijian Sun	2118
<i>Multimodal Joint Attribute Prediction and Value Extraction for E-commerce Product</i>	
Tiangang Zhu, Yue Wang, Haoran Li, Youzheng Wu, Xiaodong He and Bowen Zhou	2129
<i>A Predicate-Function-Argument Annotation of Natural Language for Open-Domain Information Expression</i>	
Mingming Sun, Wenyue Hua, Zoey Liu, Xin Wang, Kangjie Zheng and Ping Li	2140
<i>Retrofitting Structure-aware Transformer Language Model for End Tasks</i>	
Hao Fei, Yafeng Ren and Donghong Ji	2151
<i>Lightweight, Dynamic Graph Convolutional Networks for AMR-to-Text Generation</i>	
Yan Zhang, Zhijiang Guo, Zhiyang Teng, Wei Lu, Shay B. Cohen, ZUOZHU LIU and Lidong Bing	2162
<i>If beam search is the answer, what was the question?</i>	
Clara Meister, Ryan Cotterell and Tim Vieira	2173
<i>Understanding the Mechanics of SPIGOT: Surrogate Gradients for Latent Structure Learning</i>	
Tsvetomila Mihaylova, Vlad Niculae and André F. T. Martins	2186
<i>Is the best better? Bayesian statistical model comparison for natural language processing</i>	
Piotr Szymański and Kyle Gorman	2203
<i>Multi-Task Learning for Logically Dependent Tasks from the Perspective of Causal Inference</i>	
Wenqing Chen, Jidong Tian, Liqiang Xiao, Hao He and Yaohui Jin	2213
<i>Masking as an Efficient Alternative to Finetuning for Pretrained Language Models</i>	
Mengjie Zhao, Tao Lin, Fei Mi, Martin Jaggi and Hinrich Schütze	2226
<i>Dynamic Context Selection for Document-level Neural Machine Translation via Reinforcement Learning</i>	
Xiaomian Kang, Yang Zhao, Jiajun Zhang and Chengqing Zong	2242

<i>Data Rejuvenation: Exploiting Inactive Training Examples for Neural Machine Translation</i> Wenxiang Jiao, Xing Wang, Shilin He, Irwin King, Michael Lyu and Zhaopeng Tu	2255
<i>Targeted Finetuning for NMT with Conditional Generative-Discriminative Loss</i> Prathyusha Jwalapuram, Shafiq Joty and Youlin Shen	2267
<i>Learning Adaptive Segmentation Policy for Simultaneous Translation</i> Ruiqing Zhang, Chuanqiang Zhang, Zhongjun He, Hua Wu and Haifeng Wang	2280
<i>Learn to Cross-lingual Transfer with Meta Graph Learning Across Heterogeneous Languages</i> Zheng Li, Mukul Kumar, William Headden, Bing Yin, Ying Wei, Yu Zhang and Qiang Yang .	2290
<i>UDapter: Language Adaptation for Truly Universal Dependency Parsing</i> Ahmet Üstün, Arianna Bisazza, Gosse Bouma and Gertjan van Noord	2302
<i>Uncertainty-Aware Label Refinement for Sequence Labeling</i> Tao Gui, Jiacheng Ye, Qi Zhang, Zhengyan Li, Zichu Fei, Yeyun Gong and Xuanjing Huang .	2316
<i>Adversarial Attack and Defense of Structured Prediction Models</i> Wenjuan Han, Liwen Zhang, Yong Jiang and Kewei Tu	2327
<i>Position-Aware Tagging for Aspect Sentiment Triplet Extraction</i> Lu Xu, Hao Li, Wei Lu and Lidong Bing	2339
<i>Simultaneous Machine Translation with Visual Context</i> Ozan Caglayan, Julia Ive, Veneta Haralampieva, Pranava Madhyastha, Loïc Barrault and Lucia Specia	2350
<i>XCOPA: A Multilingual Dataset for Causal Commonsense Reasoning</i> Edoardo Maria Ponti, Goran Glavaš, Olga Majewska, Qianchu Liu, Ivan Vulić and Anna Korhonen	2362
<i>The Secret is in the Spectra: Predicting Cross-lingual Task Performance with Spectral Similarity Measures</i> Haim Dubossarsky, Ivan Vulić, Roi Reichart and Anna Korhonen	2377
<i>Bridging linguistic typology and multilingual machine translation with multi-view language representations</i> Arturo Oncevay, Barry Haddow and Alexandra Birch	2391
<i>AnswerFact: Fact Checking in Product Question Answering</i> Wenxuan Zhang, Yang Deng, Jing Ma and Wai Lam	2407
<i>Context-Aware Answer Extraction in Question Answering</i> Yeon Seonwoo, Ji-Hoon Kim, Jung-Woo Ha and Alice Oh	2418
<i>What do Models Learn from Question Answering Datasets?</i> Priyanka Sen and Amir Saffari	2429
<i>Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading</i> Yifan Gao, Chien-Sheng Wu, Jingjing Li, Shafiq Joty, Steven C.H. Hoi, Caiming Xiong, Irwin King and Michael Lyu	2439
<i>A Method for Building a Commonsense Inference Dataset based on Basic Events</i> Kazumasa Omura, Daisuke Kawahara and Sadao Kurohashi	2450

<i>Neural Deepfake Detection with Factual Structure of Text</i>	
Wanjun Zhong, Duyu Tang, Zenan Xu, Ruize Wang, Nan Duan, Ming Zhou, Jiahai Wang and Jian Yin	2461
<i>MultiCQA: Zero-Shot Transfer of Self-Supervised Text Matching Models on a Massive Scale</i>	
Andreas Rücklé, Jonas Pfeiffer and Iryna Gurevych	2471
<i>Enabling Cross-Lingual AMR Parsing with Transfer Learning Techniques</i>	
Rexhina Blloshmi, Rocco Tripodi and Roberto Navigli	2487
<i>Improving AMR parsing with Sequence-to-Sequence Pre-training</i>	
Dongqin Xu, Junhui Li, Muhua Zhu, Min Zhang and Guodong Zhou	2501
<i>Hate-Speech and Offensive Language Detection in Roman Urdu</i>	
Hammad Rizwan, Muhammad Haroon Shakeel and Asim Karim	2512
<i>Suicidal Risk Detection for Military Personnel</i>	
Sungjoon Park, Kiwoong Park, Jaimeen Ahn and Alice Oh	2523
<i>Comparative Evaluation of Label Agnostic Selection Bias in Multilingual Hate Speech Datasets</i>	
Nedjma Ousidhoum, Yangqiu Song and Dit-Yan Yeung	2532
<i>HENIN: Learning Heterogeneous Neural Interaction Networks for Explainable Cyberbullying Detection on Social Media</i>	
Hsin-Yu Chen and Cheng-Te Li	2543
<i>I Was Just Being Sarcastic! Reactive Supervision: A New Method for Collecting Sarcasm Data</i>	
Boaz Shmueli, Lun-Wei Ku and Soumya Ray	2553
<i>Self-Induced Curriculum Learning in Self-Supervised Neural Machine Translation</i>	
Dana Ruitter, Josef van Genabith and Cristina España-Bonet	2560
<i>Towards Reasonably-Sized Character-Level Transformer NMT by Finetuning Subword Systems</i>	
Jindřich Libovický and Alexander Fraser	2572
<i>Transfer Learning and Distant Supervision for Multilingual Transformer Models: A Study on African Languages</i>	
Michael A. Hedderich, David Adelani, Dawei Zhu, Jesujoba Alabi, Udia Markus and Dietrich Klakow	2580
<i>Translation Quality Estimation by Jointly Learning to Score and Rank</i>	
Jingyi Zhang and Josef van Genabith	2592
<i>Direct Segmentation Models for Streaming Speech Translation</i>	
Javier Iranzo-Sánchez, Adrià Giménez Pastor, Joan Albert Silvestre-Cerdà, Pau Baquero-Arnal, Jorge Civera Saiz and Alfons Juan	2599
<i>Not Low-Resource Anymore: Aligner Ensembling, Batch Filtering, and New Datasets for Bengali-English Machine Translation</i>	
Tahmid Hasan, Abhik Bhattacharjee, Kazi Samin, Masum Hasan, Madhusudan Basak, M. Sohel Rahman and Rifat Shahriyar	2612
<i>CSP: Code-Switching Pre-training for Neural Machine Translation</i>	
Zhen Yang, Bojie Hu, ambyera han, shen huang and Qi Ju	2624

<i>Type B Reflexivization as an Unambiguous Testbed for Multilingual Multi-Task Gender Bias</i>	
Ana Valeria González, Maria Barrett, Rasmus Hvingelby, Kellie Webster and Anders Søgaard	2637
<i>Pre-training Multilingual Neural Machine Translation by Leveraging Alignment Information</i>	
Zehui Lin, Xiao Pan, Mingxuan Wang, Xipeng Qiu, Jiangtao Feng, Hao Zhou and Lei Li	2649
<i>Losing Heads in the Lottery: Pruning Transformer Attention in Neural Machine Translation</i>	
Maximiliana Behnke and Kenneth Heafield	2664
<i>Towards Enhancing Faithfulness for Neural Machine Translation</i>	
Rongxiang Weng, Heng Yu, Xiangpeng Wei and Weihua Luo	2675
<i>COMET: A Neural Framework for MT Evaluation</i>	
Ricardo Rei, Craig Stewart, Ana C Farinha and Alon Lavie	2685
<i>Reusing a Pretrained Language Model on Languages with Limited Corpora for Unsupervised NMT</i>	
Alexandra Chronopoulou, Dario Stojanovski and Alexander Fraser	2703
<i>LNMap: Departures from Isomorphic Assumption in Bilingual Lexicon Induction Through Non-Linear Mapping in Latent Space</i>	
Tasnim Mohiuddin, M Saiful Bari and Shafiq Joty	2712
<i>Uncertainty-Aware Semantic Augmentation for Neural Machine Translation</i>	
Xiangpeng Wei, Heng Yu, Yue Hu, Rongxiang Weng, Luxi Xing and Weihua Luo	2724
<i>Can Automatic Post-Editing Improve NMT?</i>	
Shamil Chollampatt, Raymond Hendy Susanto, Liling Tan and Ewa Szymanska	2736
<i>Parsing Gapping Constructions Based on Grammatical and Semantic Roles</i>	
Yoshihide Kato and Shigeki Matsubara	2747
<i>Span-based discontinuous constituency parsing: a family of exact chart-based algorithms with time complexities from $O(n^6)$ down to $O(n^3)$</i>	
Caio Corro	2753
<i>Some Languages Seem Easier to Parse Because Their Treebanks Leak</i>	
Anders Søgaard	2765
<i>Discontinuous Constituent Parsing as Sequence Labeling</i>	
David Vilares and Carlos Gómez-Rodríguez	2771
<i>Modularized Syntactic Neural Networks for Sentence Classification</i>	
Haiyan Wu, Ying Liu and Shaoyun Shi	2786
<i>TED-CDB: A Large-Scale Chinese Discourse Relation Dataset on TED Talks</i>	
Wanqiu Long, Bonnie Webber and Deyi Xiong	2793
<i>QADiscourse - Discourse Relations as QA Pairs: Representation, Crowdsourcing and Baselines</i>	
Valentina Pyatkin, Ayal Klein, Reut Tsarfaty and Ido Dagan	2804
<i>Discourse Self-Attention for Discourse Element Identification in Argumentative Student Essays</i>	
Wei Song, Ziyao Song, Ruiji Fu, Lizhen Liu, Miaomiao Cheng and Ting Liu	2820
<i>Controllable Story Generation with External Knowledge Using Large-Scale Language Models</i>	
Peng Xu, Mostofa Patwary, Mohammad Shoebybi, Raul Puri, Pascale Fung, Anima Anandkumar and Bryan Catanzaro	2831

<i>Incomplete Utterance Rewriting as Semantic Segmentation</i>	
Qian Liu, Bei Chen, Jian-Guang LOU, Bin Zhou and Dongmei Zhang	2846
<i>Improving Grammatical Error Correction Models with Purpose-Built Adversarial Examples</i>	
Lihao Wang and Xiaoqing Zheng	2858
<i>Homophonic Pun Generation with Lexically Constrained Rewriting</i>	
Zhiwei Yu, Hongyu Zang and Xiaojun Wan	2870
<i>How to Make Neural Natural Language Generation as Reliable as Templates in Task-Oriented Dialogue</i>	
Henry Elder, Alexander O’Connor and Jennifer Foster	2877
<i>Multilingual AMR-to-Text Generation</i>	
Angela Fan and Claire Gardent	2889
<i>Exploring the Linear Subspace Hypothesis in Gender Bias Mitigation</i>	
Francisco Vargas and Ryan Cotterell	2902
<i>Lifelong Language Knowledge Distillation</i>	
Yung-Sung Chuang, Shang-Yu Su and Yun-Nung Chen	2914
<i>Sparse Parallel Training for Hierarchical Dirichlet Process Topic Models</i>	
Alexander Terenin, Måns Magnusson and Leif Jonsson	2925
<i>Multi-label Few/Zero-shot Learning with Knowledge Aggregated from Multiple Label Graphs</i>	
Jueqing Lu, Lan Du, Ming Liu and Joanna Dipnall	2935
<i>Word Rotator’s Distance</i>	
Sho Yokoi, Ryo Takahashi, Reina Akama, Jun Suzuki and Kentaro Inui	2944
<i>Disentangle-based Continual Graph Representation Learning</i>	
Xiaoyu Kou, Yankai Lin, Shaobo Liu, Peng Li, Jie Zhou and Yan Zhang	2961
<i>Semi-Supervised Bilingual Lexicon Induction with Two-way Interaction</i>	
Xu Zhao, Zihao Wang, Hao Wu and Yong Zhang	2973
<i>Wasserstein Distance Regularized Sequence Representation for Text Matching in Asymmetrical Domains</i>	
Weijie Yu, Chen Xu, Jun Xu, Liang Pang, Xiaopeng Gao, Xiaozhao Wang and Ji-Rong Wen .	2985
<i>A Simple Approach to Learning Unsupervised Multilingual Embeddings</i>	
Pratik Jawanpuria, Mayank Meghwanshi and Bamdev Mishra	2995
<i>Bootstrapped Q-learning with Context Relevant Observation Pruning to Generalize in Text-based Games</i>	
Subhajit Chaudhury, Daiki Kimura, Kartik Talamadupula, Michiaki Tatsubori, Asim Munawar and Ryuki Tachibana	3002
<i>BERT-EMD: Many-to-Many Layer Mapping for BERT Compression with Earth Mover’s Distance</i>	
jianquan li, Xiaokang Liu, Honghong Zhao, Ruifeng Xu, Min Yang and yaohong jin	3009
<i>Slot Attention with Value Normalization for Multi-domain Dialogue State Tracking</i>	
Yexiang Wang, Yi Guo and Siqi Zhu	3019
<i>Don’t Read Too Much Into It: Adaptive Computation for Open-Domain Question Answering</i>	
Yuxiang Wu, Sebastian Riedel, Pasquale Minervini and Pontus Stenetorp	3029

<i>Multi-Step Inference for Reasoning Over Paragraphs</i>	
Jiangming Liu, Matt Gardner, Shay B. Cohen and Mirella Lapata	3040
<i>Learning a Cost-Effective Annotation Policy for Question Answering</i>	
Bernhard Kratzwald, Stefan Feuerriegel and Huan Sun	3051
<i>Scene Restoring for Narrative Machine Reading Comprehension</i>	
Zhixing Tian, Yuanzhe Zhang, Kang Liu, Jun Zhao, Yantao Jia and Zhicheng Sheng	3063
<i>A Simple and Effective Model for Answering Multi-span Questions</i>	
Elad Segal, Avia Efrat, Mor Shoham, Amir Globerson and Jonathan Berant	3074
<i>Top-Rank-Focused Adaptive Vote Collection for the Evaluation of Domain-Specific Semantic Models</i>	
Pierangelo Lombardo, Alessio Boiardi, Luca Colombo, Angelo Schiavone and Nicolò Tamagnone	3081
<i>Meta Fine-Tuning Neural Language Models for Multi-Domain Text Mining</i>	
Chengyu Wang, Minghui Qiu, jun huang and XIAOFENG HE	3094
<i>Incorporating Context Structures for Query Generation</i>	
Ruey-Cheng Chen and Chia-Jung Lee	3105
<i>Conditional Causal Relationships between Emotions and Causes in Texts</i>	
Xinhong Chen, Qing Li and Jianping Wang	3111
<i>COMETA: A Corpus for Medical Entity Linking in the Social Media</i>	
Marco Basaldella, Fangyu Liu, Ehsan Shareghi and Nigel Collier	3122
<i>Pareto Probing: Trading-Off Accuracy and Complexity</i>	
Tiago Pimentel, Naomi Saphra, Adina Williams and Ryan Cotterell	3138
<i>Interpretation of NLP models through input marginalization</i>	
Siwon Kim, Jihun Yi, Eunji Kim and Sungroh Yoon	3154
<i>Generating Label Cohesive and Well-Formed Adversarial Claims</i>	
Pepa Atanasova, Dustin Wright and Isabelle Augenstein	3168
<i>Are All Good Word Vector Spaces Isomorphic?</i>	
Ivan Vulić, Sebastian Ruder and Anders Søgaard	3178
<i>Cold-start and Interpretability: Turning Regular Expressions into Trainable Recurrent Neural Networks</i>	
Chengyue Jiang, Yinggong Zhao, Shanbo Chu, Libin Shen and Kewei Tu	3193
<i>When BERT Plays the Lottery, All Tickets Are Winning</i>	
Sai Prasanna, Anna Rogers and Anna Rumshisky	3208
<i>On the weak link between importance and prunability of attention heads</i>	
Aakriti Budhraj, Madhura Pande, Preksha Nema, Pratyush Kumar and Mitesh M. Khapra . . .	3230
<i>Towards Interpreting BERT for Reading Comprehension Based QA</i>	
Sahana Ramnath, Preksha Nema, Deep Sahni and Mitesh M. Khapra	3236
<i>How do Decisions Emerge across Layers in Neural Models? Interpretation with Differentiable Masking</i>	
Nicola De Cao, Michael Sejr Schlichtkrull, Wilker Aziz and Ivan Titov	3243

<i>A Diagnostic Study of Explainability Techniques for Text Classification</i>	
Pepa Atanasova, Jakob Grue Simonsen, Christina Lioma and Isabelle Augenstein	3256
<i>STL-CQA: Structure-based Transformers with Localization and Encoding for Chart Question Answering</i>	
Hrituraj Singh and Sumit Shekhar	3275
<i>Learning to Contrast the Counterfactual Samples for Robust Visual Question Answering</i>	
Zujie Liang, Weitao Jiang, Haifeng Hu and Jiaying Zhu	3285
<i>Learning Physical Common Sense as Knowledge Graph Completion via BERT Data Augmentation and Constrained Tucker Factorization</i>	
Zhenjie Zhao, Evangelos Papalexakis and Xiaojuan Ma	3293
<i>A Visually-grounded First-person Dialogue Dataset with Verbal and Non-verbal Responses</i>	
Hisashi Kamezawa, Noriki Nishida, Nobuyuki Shimizu, Takashi Miyazaki and Hideki Nakayama	3299
<i>Cross-Media Keyphrase Prediction: A Unified Framework with Multi-Modality Multi-Head Attention and Image Wordings</i>	
Yue Wang, Jing Li, Michael Lyu and Irwin King	3311
<i>VD-BERT: A Unified Vision and Dialog Transformer with BERT</i>	
Yue Wang, Shafiq Joty, Michael Lyu, Irwin King, Caiming Xiong and Steven C.H. Hoi	3325
<i>The Grammar of Emergent Languages</i>	
Oskar van der Wal, Silvan de Boer, Elia Bruni and Dieuwke Hupkes	3339
<i>Sub-Instruction Aware Vision-and-Language Navigation</i>	
Yicong Hong, Cristian Rodriguez, Qi Wu and Stephen Gould	3360
<i>Knowledge-Grounded Dialogue Generation with Pre-trained Language Models</i>	
Xueliang Zhao, wei wu, Can Xu, Chongyang Tao, Dongyan Zhao and Rui Yan	3377
<i>MinTL: Minimalist Transfer Learning for Task-Oriented Dialogue Systems</i>	
Zhaojiang Lin, Andrea Madotto, Genta Indra Winata and Pascale Fung	3391
<i>Variational Hierarchical Dialog Autoencoder for Dialog State Tracking Data Augmentation</i>	
Kang Min Yoo, Hanbit Lee, Franck Dernoncourt, Trung Bui, Walter Chang and Sang-goo Lee	3406
<i>Bridging the Gap between Prior and Posterior Knowledge Selection for Knowledge-Grounded Dialogue Generation</i>	
Xiuyi Chen, Fandong Meng, Peng Li, Feilong Chen, Shuang Xu, Bo Xu and Jie Zhou	3426
<i>Counterfactual Off-Policy Training for Neural Dialogue Generation</i>	
Qingfu Zhu, Wei-Nan Zhang, Ting Liu and William Yang Wang	3438
<i>Dialogue Distillation: Open-domain Dialogue Augmentation Using Unpaired Data</i>	
Rongsheng Zhang, Yinhe Zheng, Jianzhi Shao, Xiaoxi Mao, Yadong Xi and Minlie Huang . .	3449
<i>Task-Completion Dialogue Policy Learning via Monte Carlo Tree Search with Dueling Network</i>	
Sihan Wang, kaijie zhou, Kunfeng Lai and Jianping Shen	3461
<i>Learning a Simple and Effective Model for Multi-turn Response Generation with Auxiliary Tasks</i>	
YUFAN ZHAO, Can Xu and wei wu	3472

<i>AttnIO: Knowledge Graph Exploration with In-and-Out Attention Flow for Knowledge-Grounded Dialogue</i>	
Jaehun Jung, Bokyung Son and Sungwon Lyu	3484
<i>Amalgamating Knowledge from Two Teachers for Task-oriented Dialogue System with Adversarial Training</i>	
Wanwei He, Min Yang, Rui Yan, Chengming Li, Ying Shen and Ruifeng Xu	3498
<i>Task-oriented Domain-specific Meta-Embedding for Text Classification</i>	
Xin Wu, Yi Cai, Yang Kai, Tao Wang and Qing Li	3508
<i>Don't Neglect the Obvious: On the Role of Unambiguous Words in Word Sense Disambiguation</i>	
Daniel Loureiro and Jose Camacho-Collados	3514
<i>Within-Between Lexical Relation Classification using Path-based and Distributional Data</i>	
Oren Barkan, Avi Caciularu and Ido Dagan	3521
<i>With More Contexts Comes Better Performance: Contextualized Sense Embeddings for All-Round Word Sense Disambiguation</i>	
Bianca Scarlini, Tommaso Pasini and Roberto Navigli	3528
<i>Convolution over Hierarchical Syntactic and Lexical Graphs for Aspect Level Sentiment Analysis</i>	
Mi Zhang and Tiejun Qian	3540
<i>Multi-Instance Multi-Label Learning Networks for Aspect-Category Sentiment Analysis</i>	
Yuncong Li, Cunxiang Yin, Sheng-hua Zhong and Xu Pan	3550
<i>Aspect Based Sentiment Analysis with Aspect-Specific Opinion Spans</i>	
Lu Xu, Lidong Bing, Wei Lu and Fei Huang	3561
<i>Emotion-Cause Pair Extraction as Sequence Labeling Based on A Novel Tagging Scheme</i>	
Chaofa Yuan, Chuang Fan, Jianzhu Bao and Ruifeng Xu	3568
<i>End-to-End Emotion-Cause Pair Extraction based on Sliding Window Multi-Label Learning</i>	
Zixiang Ding, Rui Xia and Jianfei Yu	3574
<i>Multi-modal Multi-label Emotion Detection with Modality and Label Dependence</i>	
Dong Zhang, Xincheng Ju, Junhui Li, Shoushan Li, Qiaoming Zhu and Guodong Zhou	3584
<i>Tasty Burgers, Soggy Fries: Probing Aspect Robustness in Aspect-Based Sentiment Analysis</i>	
Xiaoyu Xing, Zhijing Jin, Di Jin, Bingning Wang, Qi Zhang and Xuanjing Huang	3594
<i>Modeling Content Importance for Summarization with Pre-trained Language Models</i>	
Liqiang Xiao, Lu Wang, Hao He and Yaohui Jin	3606
<i>Unsupervised Reference-Free Summary Quality Evaluation via Contrastive Learning</i>	
Hanlu Wu, Tengfei Ma, Lingfei Wu, Tariro Manyumwa and Shouling Ji	3612
<i>Neural Extractive Summarization with Hierarchical Attentive Heterogeneous Graph Network</i>	
Ruipeng Jia, Yanan Cao, Hengzhu Tang, Fang Fang, Cong Cao and Shi Wang	3622
<i>Coarse-to-Fine Query Focused Multi-Document Summarization</i>	
Yumo Xu and Mirella Lapata	3632
<i>Pre-training for Abstractive Document Summarization by Reinstating Source Text</i>	
Yanyan Zou, Xingxing Zhang, Wei Lu, Furu Wei and Ming Zhou	3646

<i>Learning from Context or Names? An Empirical Study on Neural Relation Extraction</i>	
Hao Peng, Tianyu Gao, Xu Han, Yankai Lin, Peng Li, Zhiyuan Liu, Maosong Sun and Jie Zhou	
	3661
<i>SelfORE: Self-supervised Relational Feature Learning for Open Relation Extraction</i>	
Xuming Hu, Lijie Wen, Yusong Xu, Chenwei Zhang and Philip Yu	3673
<i>Denoising Relation Extraction from Document-level Distant Supervision</i>	
Chaojun Xiao, Yuan Yao, Ruobing Xie, Xu Han, Zhiyuan Liu, Maosong Sun, Fen Lin and Leyu Lin	3683
<i>Let's Stop Error Propagation in the End-to-End Relation Extraction Literature!</i>	
Bruno Taillé, Vincent Guigue, Geoffrey Scoutheeten and patrick Gallinari	3689
<i>Exposing Shallow Heuristics of Relation Extraction Models with Challenge Data</i>	
Shachar Rosenman, Alon Jacovi and Yoav Golderg	3702
<i>Global-to-Local Neural Networks for Document-Level Relation Extraction</i>	
Difeng Wang, Wei Hu, Ermei Cao and Weijian Sun	3711
<i>Recurrent Interaction Network for Jointly Extracting Entities and Classifying Relations</i>	
Kai Sun, Richong Zhang, Samuel Mensah, Yongyi Mao and xudong Liu	3722
<i>Temporal Knowledge Base Completion: New Algorithms and Evaluation Protocols</i>	
Prachi Jain, Sushant Rathi, and Soumen Chakrabarti	3733
<i>Constrained Iterative Labeling for Open Information Extraction</i>	
Keshav Kolluru, Vaibhav Adlakha, Samarth Aggarwal, and Soumen Chakrabarti	3748
<i>Public Sentiment Drift Analysis Based on Hierarchical Variational Auto-encoder</i>	
Wenyue Zhang, Xiaoli Li, Yang Li, Suge Wang, Deyu Li, Jian Liao and Jianxing Zheng	3762
<i>Point to the Expression: Solving Algebraic Word Problems using the Expression-Pointer Transformer Model</i>	
Bugeun Kim, Kyung Seo Ki, Donggeon Lee and Gahgene Gweon	3768
<i>Semantically-Aligned Universal Tree-Structured Solver for Math Word Problems</i>	
Jinghui Qin, Lihui Lin, Xiaodan Liang, Rumin Zhang and Liang Lin	3780
<i>Neural Topic Modeling by Incorporating Document Relationship Graph</i>	
Deyu Zhou, Xuemeng Hu and Rui Wang	3790
<i>Routing Enforced Generative Model for Recipe Generation</i>	
Zhiwei Yu, Hongyu Zang and Xiaojun Wan	3797
<i>Assessing the Helpfulness of Learning Materials with Inference-Based Learner-Like Agent</i>	
Yun-Hsuan Jen, Chieh-Yang Huang, MeiHua Chen, Ting-Hao Huang and Lun-Wei Ku	3807
<i>Selection and Generation: Learning towards Multi-Product Advertisement Post Generation</i>	
Zhangming Chan, Yuchi Zhang, Xiuying Chen, Shen Gao, Zhiqiang Zhang, Dongyan Zhao and Rui Yan	3818
<i>Form2Seq: A Framework for Higher-Order Form Structure Extraction</i>	
Milan Aggarwal, Hires Gupta, Mausoom Sarkar and Balaji Krishnamurthy	3830

<i>Domain Adaptation of Thai Word Segmentation Models using Stacked Ensemble</i>	
Peerat Limkonchotiwat, Wannaphong Phatthiyaphaibun, Raheem Sarwar, Ekapol Chuangsuwanich and Sarana Nutanong	3841
<i>DagoBERT: Generating Derivational Morphology with a Pretrained Language Model</i>	
Valentin Hofmann, Janet Pierrehumbert and Hinrich Schütze	3848
<i>Attention Is All You Need for Chinese Word Segmentation</i>	
Sufeng Duan and Hai Zhao	3862
<i>A Joint Multiple Criteria Model in Transfer Learning for Cross-domain Chinese Word Segmentation</i>	
Kaiyu Huang, Degen Huang, Zhuang Liu and Fengran Mo	3873
<i>Alignment-free Cross-lingual Semantic Role Labeling</i>	
Rui Cai and Mirella Lapata	3883
<i>Leveraging Declarative Knowledge in Text and First-Order Logic for Fine-Grained Propaganda Detection</i>	
Ruize Wang, Duyu Tang, Nan Duan, Wanjun Zhong, Zhongyu Wei, Xuanjing Huang, Daxin Jiang and Ming Zhou	3895
<i>X-SRL: A Parallel Cross-Lingual Semantic Role Labeling Dataset</i>	
Angel Daza and Anette Frank	3904
<i>Graph Convolutions over Constituent Trees for Syntax-Aware Semantic Role Labeling</i>	
Diego Marcheggiani and Ivan Titov	3915
<i>Fast semantic parsing with well-typedness guarantees</i>	
Matthias Lindemann, Jonas Groschwitz and Alexander Koller	3929
<i>Improving Out-of-Scope Detection in Intent Classification by Using Embeddings of the Word Graph Space of the Classes</i>	
Paulo Cavalin, Victor Henrique Alves Ribeiro, Ana Appel and Claudio Pinhanez	3952
<i>Supervised Seeded Iterated Learning for Interactive Language Learning</i>	
Yuchen Lu, Soumye Singhal, Florian Strub, Olivier Pietquin and Aaron Courville	3962
<i>Spot The Bot: A Robust and Efficient Framework for the Evaluation of Conversational Dialogue Systems</i>	
Jan Deriu, Don Tuggener, Pius von Däniken, Jon Ander Campos, Alvaro Rodrigo, Thiziri Belkacem, Aitor Soroa, Eneko Agirre and Mark Cieliebak	3971
<i>Human-centric dialog training via offline reinforcement learning</i>	
Natasha Jaques, Judy Hanwen Shen, Asma Ghandeharioun, Craig Ferguson, Agata Lapedriza, Noah Jones, Shixiang Gu and Rosalind Picard	3985
<i>Speakers Fill Semantic Gaps with Context</i>	
Tiago Pimentel, Rowan Hall Maudslay, Damian Blasi and Ryan Cotterell	4004
<i>Investigating Cross-Linguistic Adjective Ordering Tendencies with a Latent-Variable Model</i>	
Jun Yen Leung, Guy Emerson and Ryan Cotterell	4016
<i>Surprisal Predicts Code-Switching in Chinese-English Bilingual Text</i>	
Jesús Calvillo, Le Fang, Jeremy Cole and David Reitter	4029
<i>Investigating Lexical Variability in Language Models</i>	
Charles Yu, Ryan Sie, Nicolas Tedeschi and Leon Bergen	4040

<i>Improving Word Sense Disambiguation with Translations</i>	
Yixing Luan, Bradley Hauer, Lili Mou and Grzegorz Kondrak	4055
<i>Towards Better Context-aware Lexical Semantics: Adjusting Contextualized Representations through Static Anchors</i>	
Qianchu Liu, Diana McCarthy and Anna Korhonen	4066
<i>Compositional Demographic Word Embeddings</i>	
Charles Welch, Jonathan K. Kummerfeld, Verónica Pérez-Rosas and Rada Mihalcea	4076
<i>Do “Undocumented Immigrants” == “Illegal Aliens”? Differentiating Denotation and Connotation in Vector Space</i>	
Albert Webson, Zhizhong Chen, Carsten Eickhoff and Ellie Pavlick	4090
<i>Multi-View Sequence-to-Sequence Models with Conversational Structure for Abstractive Dialogue Summarization</i>	
Jiaao Chen and Diyi Yang	4106
<i>Few-Shot Learning for Opinion Summarization</i>	
Arthur Bražinskas, Mirella Lapata and Ivan Titov	4119
<i>Learning to Fuse Sentences with Transformers for Summarization</i>	
Logan Lebanoff, Franck Dernoncourt, Doo Soon Kim, Lidan Wang, Walter Chang and Fei Liu	4136
<i>Stepwise Extractive Summarization and Planning with Structured Transformers</i>	
Shashi Narayan, Joshua Maynez, Jakub Adamek, Daniele Pighin, Blaz Bratanić and Ryan McDonald	4143
<i>CLIRMatrix: A massively large collection of bilingual and multilingual datasets for Cross-Lingual Information Retrieval</i>	
Shuo Sun and Kevin Duh	4160
<i>SLEDGE: A Simple Yet Effective Zero-Shot Baseline for Coronavirus Scientific Knowledge Search</i>	
Sean MacAvaney, Arman Cohan and Nazli Goharian	4171
<i>Modularized Transformer-based Ranking Framework</i>	
Luyu Gao, Zhuyun Dai and Jamie Callan	4180
<i>Ad-hoc Document Retrieval using Weak-Supervision with BERT and GPT2</i>	
Yosi Mass and Haggai Roitman	4191
<i>Adversarial Semantic Collisions</i>	
Congzheng Song, Alexander Rush and Vitaly Shmatikov	4198
<i>Learning Explainable Linguistic Expressions with Neural Inductive Logic Programming for Sentence Classification</i>	
Prithviraj Sen, Marina Danilevsky, Yunyao Li, Siddhartha Brahma, Matthias Boehm, Laura Chiticariu and Rajasekar Krishnamurthy	4211
<i>Eliciting Knowledge from Language Models Using Automatically Generated Prompts</i>	
Taylor Shin, Yasaman Razeghi, Robert L Logan IV, Eric Wallace and Sameer Singh	4222
<i>Learning Variational Word Masks to Improve the Interpretability of Neural Text Classifiers</i>	
Hanjie Chen and Yangfeng Ji	4236

<i>Sparse Text Generation</i>	
Pedro Henrique Martins, Zita Marinho and André F. T. Martins	4252
<i>PlotMachines: Outline-Conditioned Generation with Dynamic Plot State Tracking</i>	
Hannah Rashkin, Asli Celikyilmaz, Yejin Choi and Jianfeng Gao	4274
<i>Do sequence-to-sequence VAEs learn global features of sentences?</i>	
Tom Bosc and Pascal Vincent	4296
<i>Content Planning for Neural Story Generation with Aristotelian Rescoring</i>	
Seraphina Goldfarb-Tarrant, Tuhin Chakrabarty, Ralph Weischedel and Nanyun Peng	4319
<i>Generating Dialogue Responses from a Semantic Latent Space</i>	
Wei-Jen Ko, Avik Ray, Yilin Shen and Hongxia Jin	4339
<i>Refer, Reuse, Reduce: Grounding Subsequent References in Visual and Conversational Contexts</i>	
Ece Takmaz, Mario Giulianelli, Sandro Pezzelle, Arabella Sinclair and Raquel Fernández ...	4350
<i>Visually Grounded Compound PCFGs</i>	
Yanpeng Zhao and Ivan Titov	4369
<i>ALICE: Active Learning with Contrastive Natural Language Explanations</i>	
Weixin Liang, James Zou and Zhou Yu	4380
<i>Room-Across-Room: Multilingual Vision-and-Language Navigation with Dense Spatiotemporal Grounding</i>	
Alexander Ku, Peter Anderson, Roma Patel, Eugene Ie and Jason Baldridge	4392
<i>Iterative Language-Based Image Editing via Self-Supervised Counterfactual Reasoning</i>	
Tsu-Jui Fu, Xin Wang, Scott Grafton, Miguel Eckstein and William Yang Wang	4413
<i>Identifying Elements Essential for BERT's Multilinguality</i>	
Philipp Dufter and Hinrich Schütze	4423
<i>On Negative Interference in Multilingual Language Models</i>	
Zirui Wang, Zachary C. Lipton and Yulia Tsvetkov	4438
<i>Pre-tokenization of Multi-word Expressions in Cross-lingual Word Embeddings</i>	
Naoki Otani, Satoru Ozaki, Xingyuan Zhao, Yucen Li, Micael St Johns and Lori Levin ...	4451
<i>Language Adapters for Zero Shot Neural Machine Translation</i>	
Jerin Philip, Alexandre Berard, Matthias Gallé and Laurent Besacier	4465
<i>Do explicit alignments robustly improve massively multilingual encoders?</i>	
Shijie Wu and Mark Dredze	4471
<i>From Zero to Hero: On the Limitations of Zero-Shot Language Transfer with Multilingual Transformers</i>	
Anne Lauscher, Vinit Ravishankar, Ivan Vulčić and Goran Glavaš	4483
<i>Distilling Multiple Domains for Neural Machine Translation</i>	
Anna Currey, Prashant Mathur and Georgiana Dinu	4500
<i>Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation</i>	
Nils Reimers and Iryna Gurevych	4512

<i>A Streaming Approach For Efficient Batched Beam Search</i>	
Kevin Yang, Violet Yao, John DeNero and Dan Klein	4526
<i>Improving Multilingual Models with Language-Clustered Vocabularies</i>	
Hyung Won Chung, Dan Garrette, Kiat Chuan Tan and Jason Riesa	4536
<i>Zero-Shot Cross-Lingual Transfer with Meta Learning</i>	
Farhad Nooralahzadeh, Giannis Bekoulis, Johannes Bjerva and Isabelle Augenstein	4547
<i>The Multilingual Amazon Reviews Corpus</i>	
Phillip Keung, Yichao Lu, György Szarvas and Noah A. Smith	4563
<i>GLUCOSE: GeneraLized and COntextualized Story Explanations</i>	
Nasrin Mostafazadeh, Aditya Kalyanpur, Lori Moon, David Buchanan, Lauren Berkowitz, Or Biran and Jennifer Chu-Carroll	4569
<i>Character-level Representations Still Improve Semantic Parsing in the Age of BERT</i>	
Rik van Noord, Antonio Toral and Johan Bos	4587
<i>Infusing Disease Knowledge into BERT for Health Question Answering, Medical Inference and Disease Name Recognition</i>	
Yun He, Ziwei Zhu, Yin Zhang, Qin Chen and James Caverlee	4604
<i>Unsupervised Commonsense Question Answering with Self-Talk</i>	
Vered Shwartz, Peter West, Ronan Le Bras, Chandra Bhagavatula and Yejin Choi	4615
<i>Reasoning about Goals, Steps, and Temporal Ordering with WikiHow</i>	
Li Zhang, Qing Lyu and Chris Callison-Burch	4630
<i>Structural Supervision Improves Few-Shot Learning and Syntactic Generalization in Neural Language Models</i>	
Ethan Wilcox, Peng Qian, Richard Futrell, Ryosuke Kohita, Roger Levy and Miguel Ballesteros	4640
<i>Investigating representations of verb bias in neural language models</i>	
Robert Hawkins, Takateru Yamakoshi, Thomas Griffiths and Adele Goldberg	4653
<i>Generating Image Descriptions via Sequential Cross-Modal Alignment Guided by Human Gaze</i>	
Ece Takmaz, Sandro Pezzelle, Lisa Beinborn and Raquel Fernández	4664
<i>Optimus: Organizing Sentences via Pre-trained Modeling of a Latent Space</i>	
Chunyuan Li, Xiang Gao, Yuan Li, Baolin Peng, Xiujun Li, Yizhe Zhang and Jianfeng Gao ..	4678
<i>Bio-Megatron: Larger Biomedical Domain Language Model</i>	
Hoo-Chang Shin, Yang Zhang, Evelina Bakhturina, Raul Puri, Mostofa Patwary, Mohammad Shoeybi and Raghav Mani	4700
<i>Text Segmentation by Cross Segment Attention</i>	
Michal Lukasik, Boris Dadachev, Kishore Papineni and Goncalo Simoes	4707
<i>RussianSuperGLUE: A Russian Language Understanding Evaluation Benchmark</i>	
Tatiana Shavrina, Alena Fenogenova, Emelyanov Anton, Denis Shevelev, Ekaterina Artemova, Valentin Malykh, Vladislav Mikhailov, Maria Tikhonova, Andrey Chertok and Andrey Evlampiev	4717
<i>An Empirical Study of Pre-trained Transformers for Arabic Information Extraction</i>	
Wuwei Lan, Yang Chen, Wei Xu and Alan Ritter	4727

<i>TNT: Text Normalization based Pre-training of Transformers for Content Moderation</i>	
Fei Tan, Yifan Hu, Changwei Hu, Keqian Li and Kevin Yen	4735
<i>Methods for Numeracy-Preserving Word Embeddings</i>	
Dhanasekar Sundararaman, Shijing Si, Vivek Subramanian, Guoyin Wang, Devamanyu Hazarika and Lawrence Carin	4742
<i>An Empirical Investigation of Contextualized Number Prediction</i>	
Taylor Berg-Kirkpatrick and Daniel Spokoyny	4754
<i>Modeling the Music Genre Perception across Language-Bound Cultures</i>	
Elena V. Epure, Guillaume Salha, Manuel Moussallam and Romain Hennequin	4765
<i>Joint Estimation and Analysis of Risk Behavior Ratings in Movie Scripts</i>	
Victor Martinez, Krishna Somandepalli, Yalda Tehrani-Uhls and Shrikanth Narayanan . . .	4780
<i>Keep it Surprisingly Simple: A Simple First Order Graph Based Parsing Model for Joint Morphosyntactic Parsing in Sanskrit</i>	
Amrith Krishna, Ashim Gupta, Deepak Garasangi, Pavankumar Satuluri and Pawan Goyal . . .	4791
<i>Unsupervised Parsing via Constituency Tests</i>	
Steven Cao, Nikita Kitaev and Dan Klein	4798
<i>Please Mind the Root: Decoding Arborescences for Dependency Parsing</i>	
Ran Zmigrod, Tim Vieira and Ryan Cotterell	4809
<i>Unsupervised Cross-Lingual Part-of-Speech Tagging for Truly Low-Resource Scenarios</i>	
Ramy Eskander, Smaranda Muresan and Michael Collins	4820
<i>Unsupervised Parsing with S-DIORA: Single Tree Encoding for Deep Inside-Outside Recursive Autoencoders</i>	
Andrew Drozdov, Subendhu Rongali, Yi-Pei Chen, Tim O’Gorman, Mohit Iyyer and Andrew McCallum	4832
<i>Utility is in the Eye of the User: A Critique of NLP Leaderboard Design</i>	
Kawin Ethayarajh and Dan Jurafsky	4846
<i>An Empirical Investigation Towards Efficient Multi-Domain Language Model Pre-training</i>	
Kristjan Arumae, Qing Sun and Parminder Bhatia	4854
<i>Analyzing Individual Neurons in Pre-trained Language Models</i>	
Nadir Durrani, Hassan Sajjad, Fahim Dalvi and Yonatan Belinkov	4865
<i>Dissecting Span Identification Tasks with Performance Prediction</i>	
Sean Papay, Roman Klinger and Sebastian Padó	4881
<i>Assessing Phrasal Representation and Composition in Transformers</i>	
Lang Yu and Allyson Ettinger	4896
<i>Analyzing Redundancy in Pretrained Transformer Models</i>	
Fahim Dalvi, Hassan Sajjad, Nadir Durrani and Yonatan Belinkov	4908
<i>Be More with Less: Hypergraph Attention Networks for Inductive Text Classification</i>	
Kaize Ding, Jianling Wang, Jundong Li, Dingcheng Li and Huan Liu	4927

<i>Entities as Experts: Sparse Memory Access with Entity Supervision</i>	
Thibault Févry, Livio Baldini Soares, Nicholas FitzGerald, Eunsol Choi and Tom Kwiatkowski	4937
<i>H2KGAT: Hierarchical Hyperbolic Knowledge Graph Attention Network</i>	
Shen Wang, Xiaokai Wei, Cicero Nogueira dos Santos, Zhiguo Wang, Ramesh Nallapati, Andrew Arnold, Bing Xiang and Philip S. Yu	4952
<i>Does the Objective Matter? Comparing Training Objectives for Pronoun Resolution</i>	
Yordan Yordanov, Oana-Maria Camburu, Vid Kocijan and Thomas Lukasiewicz	4963
<i>On Losses for Modern Language Models</i>	
Stéphane Aroca-Ouellette and Frank Rudzicz	4970
<i>We Can Detect Your Bias: Predicting the Political Ideology of News Articles</i>	
Ramy Baly, Giovanni Da San Martino, James Glass and Preslav Nakov	4982
<i>Semantic label smoothing for sequence to sequence problems</i>	
Michal Lukasik, Himanshu Jain, Aditya Menon, Seungyeon Kim, Srinadh Bhojanapalli, Felix Yu and Sanjiv Kumar	4992
<i>Training for Gibbs Sampling on Conditional Random Fields with Neural Scoring Factors</i>	
Sida Gao and Matthew R. Gormley	4999
<i>Multilevel Text Alignment with Cross-Document Attention</i>	
Xuhui Zhou, Nikolaos Pappas and Noah A. Smith	5012
<i>Conversational Semantic Parsing</i>	
Armen Aghajanyan, Jean Maillard, Akshat Shrivastava, Keith Diedrick, Michael Haeger, Haoran Li, Yashar Mehdad, Veselin Stoyanov, Anuj Kumar, Mike Lewis and Sonal Gupta	5026
<i>Probing Task-Oriented Dialogue Representation from Language Models</i>	
Chien-Sheng Wu and Caiming Xiong	5036
<i>End-to-End Slot Alignment and Recognition for Cross-Lingual NLU</i>	
Weijia Xu, Batoool Haider and Saab Mansour	5052
<i>Discriminative Nearest Neighbor Few-Shot Intent Detection by Transferring Natural Language Inference</i>	
Jianguo Zhang, Kazuma Hashimoto, Wenhao Liu, Chien-Sheng Wu, Yao Wan, Philip Yu, Richard Socher and Caiming Xiong	5064
<i>Simple Data Augmentation with the Mask Token Improves Domain Adaptation for Dialog Act Tagging</i>	
Semih Yavuz, Kazuma Hashimoto, Wenhao Liu, Nitish Shirish Keskar, Richard Socher and Caiming Xiong	5083
<i>Low-Resource Domain Adaptation for Compositional Task-Oriented Semantic Parsing</i>	
Xilun Chen, Asish Ghoshal, Yashar Mehdad, Luke Zettlemoyer and Sonal Gupta	5090
<i>Sound Natural: Content Rephrasing in Dialog Systems</i>	
Arash Einolghozati, Anchit Gupta, Keith Diedrick and Sonal Gupta	5101
<i>Zero-Shot Crosslingual Sentence Simplification</i>	
Jonathan Mallinson, Rico Sennrich and Mirella Lapata	5109
<i>Facilitating the Communication of Politeness through Fine-Grained Paraphrasing</i>	
Liye Fu, Susan Fussell and Cristian Danescu-Niculescu-Mizil	5127

<i>CAT-Gen: Improving Robustness in NLP Models via Controlled Adversarial Text Generation</i>	
Tianlu Wang, Xuezhi Wang, Yao Qin, Ben Packer, Kang Li, Jilin Chen, Alex Beutel and Ed Chi	
	5141
<i>Seq2Edits: Sequence Transduction Using Span-level Edit Operations</i>	
Felix Stahlberg and Shankar Kumar	5147
<i>Controllable Meaning Representation to Text Generation: Linearization and Data Augmentation Strategies</i>	
Chris Kedzie and Kathleen McKeown	5160
<i>Blank Language Models</i>	
Tianxiao Shen, Victor Quach, Regina Barzilay and Tommi Jaakkola	5186
<i>COD3S: Diverse Generation with Discrete Semantic Signatures</i>	
Nathaniel Weir, João Sedoc and Benjamin Van Durme	5199
<i>Automatic Extraction of Rules Governing Morphological Agreement</i>	
Aditi Chaudhary, Antonios Anastasopoulos, Adithya Pratapa, David R. Mortensen, Zaid Sheikh, Yulia Tsvetkov and Graham Neubig	5212
<i>Tackling the Low-resource Challenge for Canonical Segmentation</i>	
Manuel Mager, Özlem Çetinoğlu and Katharina Kann	5237
<i>IGT2P: From Interlinear Glossed Texts to Paradigms</i>	
Sarah Moeller, Ling Liu, Changbing Yang, Katharina Kann and Mans Hulden	5251
<i>A Computational Approach to Understanding Empathy Expressed in Text-Based Mental Health Support</i>	
Ashish Sharma, Adam Miner, David Atkins and Tim Althoff	5263
<i>Modeling Protagonist Emotions for Emotion-Aware Storytelling</i>	
Faeze Brahman and Snigdha Chaturvedi	5277
<i>Help! Need Advice on Identifying Advice</i>	
Venkata Subrahmanyam Govindarajan, Benjamin Chen, Rebecca Warholc, Katrin Erk and Junyi Jessy Li	5295
<i>Quantifying Intimacy In Language</i>	
Jiaxin Pei and David Jurgens	5307
<i>Writing Strategies for Science Communication: Data and Computational Analysis</i>	
Tal August, Lauren Kim, Katharina Reinecke and Noah A. Smith	5327
<i>Weakly Supervised Subevent Knowledge Acquisition</i>	
Wenlin Yao, Zeyu Dai, Maitreyi Ramaswamy, Bonan Min and Ruihong Huang	5345
<i>Biomedical Event Extraction as Sequence Labeling</i>	
Alan Ramponi, Rob van der Goot, Rosario Lombardo and Barbara Plank	5357
<i>Annotating Temporal Dependency Graphs via Crowdsourcing</i>	
Jiarui Yao, Haoling Qiu, Bonan Min and Nianwen Xue	5368
<i>Introducing a New Dataset for Event Detection in Cybersecurity Texts</i>	
Hieu Man Duc Trong, Duc Trong Le, Amir Pouran Ben Veyseh, Thuat Nguyen and Thien Huu Nguyen	5381

<i>CHARM: Inferring Personal Attributes from Conversations</i> Anna Tiginova, Andrew Yates, Paramita Mirza and Gerhard Weikum	5391
<i>Event Detection: Gate Diversity and Syntactic Importance Scores for Graph Convolution Neural Networks</i> Viet Dac Lai, Tuan Ngo Nguyen and Thien Huu Nguyen	5405
<i>Severing the Edge Between Before and After: Neural Architectures for Temporal Ordering of Events</i> Miguel Ballesteros, Rishita Anubhai, Shuai Wang, Nima Pourdamghani, Yogarshi Vyas, Jie Ma, Parminder Bhatia, Kathleen McKeown and Yaser Al-Onaizan	5412
<i>How Much Knowledge Can You Pack Into the Parameters of a Language Model?</i> Adam Roberts, Colin Raffel and Noam Shazeer	5418
<i>EXAMS: A Multi-subject High School Examinations Dataset for Cross-lingual and Multilingual Question Answering</i> Momchil Hardalov, Todor Mihaylov, Dimitrina Zlatkova, Yoan Dinkov, Ivan Koychev and Preslav Nakov	5427
<i>End-to-End Synthetic Data Generation for Domain Adaptation of Question Answering Systems</i> Siamak Shakeri, Cicero Nogueira dos Santos, Henghui Zhu, Patrick Ng, Feng Nan, Zhiguo Wang, Ramesh Nallapati and Bing Xiang	5445
<i>Multi-Stage Pretraining for Low-Resource Domain Adaptation</i> Rong Zhang, Revanth Gangi Reddy, Md Arifat Sultan, Vittorio Castelli, Anthony Ferritto, Radu Florian, Efsun Sarioglu Kayi, Salim Roukos, Avi Sil and Todd Ward	5461
<i>ISAAQ - Mastering Textbook Questions with Pre-trained Transformers and Bottom-Up and Top-Down Attention</i> Jose Manuel Gomez-Perez and Raúl Ortega	5469
<i>SubjQA: A Dataset for Subjectivity and Review Comprehension</i> Johannes Bjerva, Nikita Bhutani, Behzad Golshan, Wang-Chiew Tan and Isabelle Augenstein	5480
<i>Widget Captioning: Generating Natural Language Description for Mobile User Interface Elements</i> Yang Li, Gang Li, Luheng He, Jingjie Zheng, Hong Li and Zhiwei Guan	5495
<i>Unsupervised Natural Language Inference via Decoupled Multimodal Contrastive Learning</i> Wanyun Cui, Guangyu Zheng and Wei Wang	5511
<i>Digital Voicing of Silent Speech</i> David Gaddy and Dan Klein	5521
<i>Imitation Attacks and Defenses for Black-box Machine Translation Systems</i> Eric Wallace, Mitchell Stern and Dawn Song	5531
<i>Sequence-level Mixed Sample Data Augmentation</i> Demi Guo, Yoon Kim and Alexander Rush	5547
<i>Consistency of a Recurrent Language Model With Respect to Incomplete Decoding</i> Sean Welleck, Iliia Kulikov, Jaedeok Kim, Richard Yuanzhe Pang and Kyunghyun Cho	5553
<i>An Exploration of Arbitrary-Order Sequence Labeling via Energy-Based Inference Networks</i> Lifu Tu, Tianyu Liu and Kevin Gimpel	5569

<i>Ensemble Distillation for Structured Prediction: Calibrated, Accurate, Fast—Choose Three</i> Steven Reich, David Mueller and Nicholas Andrews	5583
<i>Inducing Target-specific Latent Structures for Aspect Sentiment Classification</i> Chenhua Chen, Zhiyang Teng and Yue Zhang	5596
<i>Affective Event Classification with Discourse-enhanced Self-training</i> Yuan Zhuang, Tianyu Jiang and Ellen Riloff	5608
<i>Deep Weighted MaxSAT for Aspect-based Opinion Extraction</i> Meixi Wu, Wenya Wang and Sinno Jialin Pan	5618
<i>Multi-view Story Characterization from Movie Plot Synopses and Reviews</i> Sudipta Kar, Gustavo Aguilar, Mirella Lapata and Thamar Solorio	5629
<i>Mind Your Inflections! Improving NLP for Non-Standard Englishes with Base-Inflection Encoding</i> Samson Tan, Shafiq Joty, Lav Varshney and Min-Yen Kan	5647
<i>Measuring the Similarity of Grammatical Gender Systems by Comparing Partitions</i> Arya D. McCarthy, Adina Williams, Shijia Liu, David Yarowsky and Ryan Cotterell	5664
<i>Is Chinese Word Segmentation a Solved Task? Rethinking Neural Chinese Word Segmentation</i> Jinlan Fu, Pengfei Liu, Qi Zhang and Xuanjing Huang	5676
<i>Learning to Pronounce Chinese Without a Pronunciation Dictionary</i> Christopher Chu, Scot Fang and Kevin Knight	5687
<i>Dynamic Anticipation and Completion for Multi-Hop Reasoning over Sparse Knowledge Graph</i> Xin Lv, Xu Han, Lei Hou, Juanzi Li, Zhiyuan Liu, Wei Zhang, YICHI ZHANG, Hao Kong and Suhui Wu	5694
<i>Knowledge Association with Hyperbolic Knowledge Graph Embeddings</i> Zequn Sun, Muhao Chen, Wei Hu, Chengming Wang, Jian Dai and Wei Zhang	5704
<i>Domain Knowledge Empowered Structured Neural Net for End-to-End Event Temporal Relation Extrac- tion</i> Rujun Han, Yichao Zhou and Nanyun Peng	5717
<i>TeMP: Temporal Message Passing for Temporal Knowledge Graph Completion</i> Jiapeng Wu, Meng Cao, Jackie Chi Kit Cheung and William L. Hamilton	5730
<i>Understanding the Difficulty of Training Transformers</i> Liyuan Liu, Xiaodong Liu, Jianfeng Gao, Weizhu Chen and Jiawei Han	5747
<i>An Empirical Study of Generation Order for Machine Translation</i> William Chan, Mitchell Stern, Jamie Kiros and Jakob Uszkoreit	5764
<i>Inference Strategies for Sequence Generation with Conditional Masking</i> Julia Kreutzer, George Foster and Colin Cherry	5774
<i>AmbigQA: Answering Ambiguous Open-domain Questions</i> Sewon Min, Julian Michael, Hannaneh Hajishirzi and Luke Zettlemoyer	5783

<i>Tell Me How to Ask Again: Question Data Augmentation with Controllable Rewriting in Continuous Space</i>	
Dayiheng Liu, Yeyun Gong, Jie Fu, Yu Yan, Jiusheng Chen, Jiancheng Lv, Nan Duan and Ming Zhou	5798
<i>Training Question Answering Models From Synthetic Data</i>	
Raul Puri, Ryan Spring, Mohammad Shoeybi, Mostofa Patwary and Bryan Catanzaro	5811
<i>Few-shot Complex Knowledge Base Question Answering via Meta Reinforcement Learning</i>	
Yuncheng Hua, Yuan-Fang Li, Gholamreza Haffari, Guilin Qi and Tongtong Wu	5827
<i>Multilingual Offensive Language Identification with Cross-lingual Embeddings</i>	
Tharindu Ranasinghe and Marcos Zampieri	5838
<i>Solving Historical Dictionary Codes with a Neural Language Model</i>	
Christopher Chu, Raphael Valenti and Kevin Knight	5845
<i>Beyond Geolocation: Micro-Dialect Identification in Diaglossic and Code-Switched Environments</i>	
Muhammad Abdul-Mageed, Chiyu Zhang, AbdelRahim Elmadany and Lyle Ungar	5855
<i>Dats Wassup!/: Investigating African-American Vernacular English in Transformer-Based Text Generation</i>	
Sophie Groenwold, Lily Ou, Aesha Parekh, Samhita Honnavalli, Sharon Levy, Diba Mirza and William Yang Wang	5877
<i>Iterative Domain-Repaired Back-Translation</i>	
Hao-Ran Wei, Zhirui Zhang, Boxing Chen and Weihua Luo	5884
<i>Dynamic Data Selection and Weighting for Iterative Back-Translation</i>	
Zi-Yi Dou, Antonios Anastasopoulos and Graham Neubig	5894
<i>Revisiting modularized multilingual NMT to meet industrial demands</i>	
Sungwon Lyu, Bokyoung Son, Kichang Yang and Jaekyoung Bae	5905
<i>LARQA: Language-agnostic answer retrieval from a multilingual pool</i>	
Uma Roy, Noah Constant, Rami Al-Rfou, Aditya Barua, Aaron Phillips and Yinfei Yang	5919
<i>OCR Post-Correction for Endangered Language Texts</i>	
Shruti Rijhwani, Antonios Anastasopoulos and Graham Neubig	5931
<i>X-FACTR: Multilingual Factual Knowledge Retrieval from Pretrained Language Models</i>	
Zhengbao Jiang, Antonios Anastasopoulos, Jun Araki, Haibo Ding and Graham Neubig	5943
<i>A Massive Collection of Cross-Lingual Web-Document Pairs</i>	
Ahmed El-Kishky, Vishrav Chaudhary, Francisco Guzmán and Philipp Koehn	5960
<i>Localizing Q&A Semantic Parsers for Any Language In a Day</i>	
Mehrad Moradshahi, Giovanni Campagna, Sina Semnani, Silei Xu and Monica Lam	5970
<i>Interactive Refinement of Cross-Lingual Word Embeddings</i>	
Michelle Yuan, Mozhi Zhang, Benjamin Van Durme, Leah Findlater and Jordan Boyd-Graber	5984
<i>Exploiting Sentence Order in Document Alignment</i>	
Brian Thompson and Philipp Koehn	5997

<i>XGLUE: A New Benchmark Dataset for Cross-lingual Pre-training, Understanding and Generation</i>	
Yaobo Liang, Nan Duan, Yeyun Gong, Ning Wu, Fenfei Guo, Weizhen Qi, Ming Gong, Linjun Shou, Daxin Jiang, Guihong Cao, Xiaodong Fan, Ruofei Zhang, Rahul Agrawal, Edward Cui, Sining Wei, Taroon Bharti, Ying Qiao, Jiun-Hung Chen, Winnie Wu, Shuguang Liu, Fan Yang, Daniel Campos, Rangan Majumder and Ming Zhou	6008
<i>AIN: Fast and Accurate Sequence Labeling with Approximate Inference Network</i>	
Xinyu Wang, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	6019
<i>HIT: Nested Named Entity Recognition via Head-Tail Pair and Token Interaction</i>	
Yu Wang, Yun Li, Hanghang Tong and Ziyue Zhu	6027
<i>Supertagging Combinatory Categorical Grammar with Attentive Graph Convolutional Networks</i>	
Yuanhe Tian, Yan Song and Fei Xia	6037
<i>An Effective Data Augmentation Method for Low-resource Tagging Tasks</i>	
BOSHENG DING, Linlin Liu, Lidong Bing, Canasai Kruengkrai, Thien Hai Nguyen, Shafiq Joty, Luo Si and Chunyan Miao	6045
<i>Interpretable Multi-dataset Evaluation for Named Entity Recognition</i>	
Jinlan Fu, Pengfei Liu and Graham Neubig	6058
<i>Adversarial Semantic Decoupling for Recognizing Open-Vocabulary Slots</i>	
Yuanmeng Yan, Keqing He, Hong Xu, Sihong Liu, Fanyu Meng, Min Hu and Weiran XU . . .	6070
<i>Plug and Play Autoencoders for Conditional Text Generation</i>	
Florian Mai, Nikolaos Pappas, Ivan Montero, Noah A. Smith and James Henderson	6076
<i>Structure Aware Negative Sampling in Knowledge Graphs</i>	
Kian Ahrabian, Aarash Feizi, Yasmin Salehi, William L. Hamilton and Avishek Joey Bose . . .	6093
<i>Neural Mask Generator: Learning to Generate Adaptive Word Maskings for Language Model Adaptation</i>	
Minki Kang, Moon-su Han and Sung Ju Hwang	6102
<i>Autoregressive Knowledge Distillation through Imitation Learning</i>	
Alexander Lin, Jeremy Wohlwend, Howard Chen and Tao Lei	6121
<i>T3: Tree-Autoencoder Regularized Adversarial Text Generation for Targeted Attack</i>	
Boxin Wang, Hengzhi Pei, Boyuan Pan, Qian Chen, Shuohang Wang and Bo Li	6134
<i>Structured Pruning of Large Language Models</i>	
Ziheng Wang, Jeremy Wohlwend and Tao Lei	6151
<i>Effective Unsupervised Domain Adaptation with Adversarially Trained Language Models</i>	
Thuy-Trang Vu, Dinh Phung and Gholamreza Haffari	6163
<i>BAE: BERT-based Adversarial Examples for Text Classification</i>	
Siddhant Garg and Goutham Ramakrishnan	6174
<i>Adversarial Self-Supervised Data Free Distillation for Text Classification</i>	
Xinyin Ma, Yongliang Shen, Gongfan Fang, Chen Chen, Chenghao Jia and Weiming Lu	6182
<i>BERT-ATTACK: Adversarial Attack Against BERT Using BERT</i>	
Linyang Li, Ruotian Ma, Qipeng Guo, Xiangyang Xue and Xipeng Qiu	6193

<i>The Thieves on Sesame Street are Polyglots — Extracting Multilingual Models from Monolingual APIs</i> Nitish Shirish Keskar, Bryan McCann, Caiming Xiong and Richard Socher	6203
<i>When Hearst Is not Enough: Improving Hypernymy Detection from Corpus with Distributional Models</i> Changlong Yu, Jialong Han, Peifeng Wang, Yangqiu Song, Hongming Zhang, Wilfred Ng and Shuming Shi	6208
<i>Interpreting Open-Domain Modifiers: Decomposition of Wikipedia Categories into Disambiguated Property-Value Pairs</i> Marius Pasca	6218
<i>A Synset Relation-enhanced Framework with a Try-again Mechanism for Word Sense Disambiguation</i> Ming Wang and Yinglin Wang	6229
<i>Diverse, Controllable, and Keyphrase-Aware: A Corpus and Method for News Multi-Headline Generation</i> Dayiheng Liu, Yeyun Gong, Yu Yan, Jie Fu, Bo Shao, Daxin Jiang, Jiancheng Lv and Nan Duan	6241
<i>Factual Error Correction for Abstractive Summarization Models</i> Meng Cao, Yue Dong, Jiapeng Wu and Jackie Chi Kit Cheung	6251
<i>Compressive Summarization with Plausibility and Saliency Modeling</i> Shrey Desai, Jiacheng Xu and Greg Durrett	6259
<i>Understanding Neural Abstractive Summarization Models via Uncertainty</i> Jiacheng Xu, Shrey Desai and Greg Durrett	6275
<i>Better Highlighting: Creating Sub-Sentence Summary Highlights</i> Sangwoo Cho, Kaiqiang Song, Chen Li, Dong Yu, Hassan Foroosh and Fei Liu	6282
<i>Summarizing Text on Any Aspects: A Knowledge-Informed Weakly-Supervised Approach</i> Bowen Tan, Lianhui Qin, Eric Xing and Zhiting Hu	6301
<i>BERT-enhanced Relational Sentence Ordering Network</i> Baiyun Cui, Yingming Li and Zhongfei Zhang	6310
<i>Online Conversation Disentanglement with Pointer Networks</i> Tao Yu and Shafiq Joty	6321
<i>VCDM: Leveraging Variational Bi-encoding and Deep Contextualized Word Representations for Improved Definition Modeling</i> Machel Reid, Edison Marrese-Taylor and Yutaka Matsuo	6331
<i>Coarse-to-Fine Pre-training for Named Entity Recognition</i> Xue Mengge, Bowen Yu, Zhenyu Zhang, Tingwen Liu, Yue Zhang and Bin Wang	6345
<i>Exploring and Evaluating Attributes, Values, and Structure for Entity Alignment</i> Zhiyuan Liu, Yixin Cao, Liangming Pan, Juanzi Li, Zhiyuan Liu and Tat-Seng Chua	6355
<i>Frustratingly Simple Few-Shot Named Entity Recognition with Structured Nearest Neighbor Learning</i> Yi Yang and Arzoo Katiyar	6365
<i>Learning Structured Representations of Entity Names using Active Learning and Weak Supervision</i> Kun Qian, Poornima Chozhiyath Raman, Yunyao Li and Lucian Popa	6376

<i>Entity Enhanced BERT Pre-training for Chinese NER</i>	
Chen Jia, Yuefeng Shi, Qinrong Yang and Yue Zhang	6384
<i>Scalable Zero-shot Entity Linking with Dense Entity Retrieval</i>	
Ledell Wu, Fabio Petroni, Martin Josifoski, Sebastian Riedel and Luke Zettlemoyer	6397
<i>A Dataset for Tracking Entities in Open Domain Procedural Text</i>	
Niket Tandon, Keisuke Sakaguchi, Bhavana Dalvi, Dheeraj Rajagopal, Peter Clark, Michal Guerquin, Kyle Richardson and Eduard Hovy	6408
<i>Design Challenges in Low-resource Cross-lingual Entity Linking</i>	
Xingyu Fu, Weijia Shi, Xiaodong Yu, Zian Zhao and Dan Roth	6418
<i>Efficient One-Pass End-to-End Entity Linking for Questions</i>	
Belinda Z. Li, Sewon Min, Srinivasan Iyer, Yashar Mehdad and Wen-tau Yih	6433
<i>LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention</i>	
Ikuya Yamada, Akari Asai, Hiroyuki Shindo, Hideaki Takeda and Yuji Matsumoto	6442
<i>Generating similes effortlessly like a Pro: A Style Transfer Approach for Simile Generation</i>	
Tuhin Chakrabarty, Smaranda Muresan and Nanyun Peng	6455
<i>HUSH: A Dataset and Platform for Human-in-the-Loop Story Generation</i>	
Nader Akoury, Shufan Wang, Josh Whiting, Stephen Hood, Nanyun Peng and Mohit Iyyer ..	6470
<i>Substance over Style: Document-Level Targeted Content Transfer</i>	
Allison Hegel, Sudha Rao, Asli Celikyilmaz and Bill Dolan	6485
<i>Template Guided Text Generation for Task Oriented Dialogue</i>	
Mihir Kale and Abhinav Rastogi	6505
<i>MOCHA: A Dataset for Training and Evaluating Generative Reading Comprehension Metrics</i>	
Anthony Chen, Gabriel Stanovsky, Sameer Singh and Matt Gardner	6521
<i>Self-Supervised Text Planning for Paragraph Completion Task</i>	
Dongyeop Kang and Eduard Hovy	6533
<i>Inquisitive Question Generation for High Level Text Comprehension</i>	
Wei-Jen Ko, TE-YUAN CHEN, Yiyang Huang, Greg Durrett and Junyi Jessy Li	6544
<i>Towards Persona-Based Empathetic Conversational Models</i>	
Peixiang Zhong, Chen Zhang, Hao Wang, Yong Liu and Chunyan Miao	6556
<i>Personal Information Leakage Detection in Conversations</i>	
Qiongfai Xu, Lizhen Qu, Zeyu Gao and Gholamreza Haffari	6567
<i>Response Selection for Multi-Party Conversations with Dynamic Topic Tracking</i>	
Weishi Wang, Steven C.H. Hoi and Shafiq Joty	6581
<i>Regularizing Dialogue Generation by Imitating Implicit Scenarios</i>	
Shaoyong Feng, Xuancheng Ren, Hongshen Chen, Bin Sun, Kan Li and Xu SUN	6592
<i>MovieChats: Chat like Humans in a Closed Domain</i>	
Hui Su, Xiaoyu Shen, Zhou Xiao, Zheng Zhang, Ernie Chang, Cheng Zhang, Cheng Niu and Jie Zhou	6605

<i>Conundrums in Entity Reference Resolution</i>	
Jing Lu and Vincent Ng	6620
<i>Semantic Role Labeling Guided Multi-turn Dialogue ReWriter</i>	
Kun Xu, Haochen Tan, Linfeng Song, Han Wu, Haisong Zhang, Linqi Song and Dong Yu ...	6632
<i>Continuity of Topic, Interaction, and Query: Learning to Quote in Online Conversations</i>	
Lingzhi Wang, Jing Li, Xingshan Zeng, Haisong Zhang and Kam-Fai Wong.....	6640
<i>Profile Consistency Identification for Open-domain Dialogue Agents</i>	
Haoyu Song, Yan Wang, Wei-Nan Zhang, Zhengyu Zhao, Ting Liu and Xiaojiang Liu.....	6651
<i>An Element-aware Multi-representation Model for Law Article Prediction</i>	
Huilin Zhong, Junsheng Zhou, Weiguang QU, Yunfei Long and Yanhui Gu.....	6663
<i>Recurrent Event Network: Autoregressive Structure Inference over Temporal Knowledge Graphs</i>	
Woojeong Jin, Meng Qu, Xisen Jin and Xiang Ren.....	6669
<i>Multi-resolution Annotations for Emoji Prediction</i>	
Weicheng Ma, Ruibo Liu, Lili Wang and Soroush Vosoughi	6684
<i>Less is More: Attention Supervision with Counterfactuals for Text Classification</i>	
Seungtaek Choi, Haeju Park, Jinyoung Yeo and Seung-won Hwang.....	6695
<i>MODE-LSTM: A Parameter-efficient Recurrent Network with Multi-Scale for Sentence Classification</i>	
Qianli Ma, Zhenxi Lin, Jiangyue Yan, Zipeng Chen and Lihong Yu.....	6705
<i>MSCNN: A Monomeric-Siamese Convolutional Neural Network for Extremely Imbalanced Multi-label Text Classification</i>	
Wenshuo Yang, Jiyi Li, Fumiyo Fukumoto and Yanming Ye	6716
<i>Multi-Stage Pre-training for Automated Chinese Essay Scoring</i>	
Wei Song, Kai Zhang, Ruiji Fu, Lizhen Liu, Ting Liu and Miaomiao Cheng.....	6723
<i>Multi-hop Inference for Question-driven Summarization</i>	
Yang Deng, Wenxuan Zhang and Wai Lam	6734
<i>Towards Interpretable Reasoning over Paragraph Effects in Situation</i>	
Mucheng Ren, Xiubo Geng, Tao QIN, Heyan Huang and Daxin Jiang	6745
<i>Question Directed Graph Attention Network for Numerical Reasoning over Text</i>	
Kunlong Chen, Weidi Xu, Xingyi Cheng, Zou Xiaochuan, Yuyu Zhang, Le Song, Taifeng Wang, Yuan Qi and Wei Chu.....	6759
<i>Dense Passage Retrieval for Open-Domain Question Answering</i>	
Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen and Wen-tau Yih.....	6769
<i>Distilling Structured Knowledge for Text-Based Relational Reasoning</i>	
Jin Dong, Marc-Antoine Rondeau and William L. Hamilton	6782
<i>Asking without Telling: Exploring Latent Ontologies in Contextual Representations</i>	
Julian Michael, Jan A. Botha and Ian Tenney	6792
<i>Pretrained Language Model Embryology: The Birth of ALBERT</i>	
Cheng-Han Chiang, Sung-Feng Huang and Hung-yi Lee	6813

<i>Learning Music Helps You Read: Using Transfer to Study Linguistic Structure in Language Models</i> Isabel Papadimitriou and Dan Jurafsky	6829
<i>What Do Position Embeddings Learn? An Empirical Study of Pre-Trained Language Model Positional Encoding</i> Yu-An Wang and Yun-Nung Chen.....	6840
<i>“You are grounded!”: Latent Name Artifacts in Pre-trained Language Models</i> Vered Shwartz, Rachel Rudinger and Oyvind Tafjord.....	6850
<i>Birds have four legs?! NumerSense: Probing Numerical Commonsense Knowledge of Pre-trained Language Models</i> Bill Yuchen Lin, Seyeon Lee, Rahul Khanna and Xiang Ren	6862
<i>Grounded Adaptation for Zero-shot Executable Semantic Parsing</i> Victor Zhong, Mike Lewis, Sida I. Wang and Luke Zettlemoyer	6869
<i>An Imitation Game for Learning Semantic Parsers from User Interaction</i> Ziyu Yao, Yiqi Tang, Wen-tau Yih, Huan Sun and Yu Su.....	6883
<i>IGSQL: Database Schema Interaction Graph Based Neural Model for Context-Dependent Text-to-SQL Generation</i> Yitao Cai and Xiaojun Wan.....	6903
<i>“What Do You Mean by That?” - A Parser-Independent Interactive Approach for Enhancing Text-to-SQL</i> Yuntao Li, Bei Chen, Qian Liu, Yan Gao, Jian-Guang LOU, Yan Zhang and Dongmei Zhang.	6913
<i>ChiTeSQL: A Large-Scale and Pragmatic Chinese Text-to-SQL Dataset</i> Lijie Wang, Ao Zhang, Kun Wu, Ke Sun, Zhenghua Li, Hua Wu, Min Zhang and Haifeng Wang	6923
<i>Mention Extraction and Linking for SQL Query Generation</i> Jianqiang Ma, ZEYU YAN, Shuai Pang, Yang Zhang and Jianping Shen	6936
<i>Re-examining the Role of Schema Linking in Text-to-SQL</i> Wenqiang Lei, Weixin Wang, Zhixin MA, Tian Gan, Wei Lu, Min-Yen Kan and Tat-Seng Chua	6943
<i>A Multi-Task Incremental Learning Framework with Category Name Embedding for Aspect-Category Sentiment Analysis</i> Zehui Dai, Cheng Peng, Huajie Chen and Yadong Ding	6955
<i>Train No Evil: Selective Masking for Task-guided Pre-training</i> Yuxian Gu, Zhengyan Zhang, Xiaozhi Wang, Zhiyuan Liu and Maosong Sun.....	6966
<i>SentiLARE: Linguistic Knowledge Enhanced Language Representation for Sentiment Analysis</i> Pei Ke, Haozhe Ji, Siyang Liu, Xiaoyan Zhu and Minlie Huang	6975
<i>Aspect-Based Sentiment Analysis by Aspect-Sentiment Joint Embedding</i> Jiaxin Huang, Yu Meng, Fang Guo, Heng Ji and Jiawei Han	6989
<i>Argument Pair Extraction from Peer Review and Rebuttal via Multi-task Learning</i> Liying Cheng, Lidong Bing, Qian Yu, Wei Lu and Luo Si	7000
<i>Diversified Multiple Instance Learning for Document-Level Multi-Aspect Sentiment Classification</i> Yunjie Ji, Hao Liu, Bolei He, Xinyan Xiao, Hua Wu and Yanhua Yu	7012

<i>An Empirical Study of Hyperbole</i>	
Li Kong, Chuanyi Li, Jidong Ge, Bin Luo and Vincent Ng	7024
<i>Unified Feature and Instance Based Domain adaptation for End-to-End Aspect-based Sentiment Analysis</i>	
Chenggong Gong, Jianfei Yu and Rui Xia	7035
<i>Compositional and Lexical Semantics in RoBERTa, BERT and DistilBERT: A Case Study on CoQA</i>	
Ieva Staliūnaitė and Ignacio Iacobacci	7046
<i>Attention is Not Only a Weight: Analyzing Transformers with Vector Norms</i>	
Goro Kobayashi, Tatsuki Kuribayashi, Sho Yokoi and Kentaro Inui	7057
<i>F1 is Not Enough! Models and Evaluation Towards User-Centered Explainable Question Answering</i>	
Hendrik Schuff, Heike Adel and Ngoc Thang Vu	7076
<i>On the Ability of Self-Attention Networks to Recognize Counter Languages</i>	
Satwik Bhattamishra, Kabir Ahuja and Navin Goyal	7096
<i>An Unsupervised Joint System for Text Generation from Knowledge Graphs and Semantic Parsing</i>	
Martin Schmitt, Sahand Sharifzadeh, Volker Tresp and Hinrich Schütze	7117
<i>A Dual-generator Network for Text Style Transfer Applications</i>	
Xiao Li, Guanyi Chen, Chenghua Lin and Ruizhe Li	7131
<i>A Knowledge-Aware Sequence-to-Tree Network for Math Word Problem Solving</i>	
Qinzhuo Wu, Qi Zhang, Jinlan Fu and Xuanjing Huang	7137
<i>Generating Fact Checking Briefs</i>	
Angela Fan, Aleksandra Piktus, Fabio Petroni, Guillaume Wenzek, Marzieh Saeidi, Andreas Vlachos, Antoine Bordes and Sebastian Riedel	7147
<i>Improving the Efficiency of Grammatical Error Correction with Erroneous Span Detection and Correction</i>	
Mengyun Chen, Tao Ge, Xingxing Zhang, Furu Wei and Ming Zhou	7162
<i>Coreferential Reasoning Learning for Language Representation</i>	
Deming Ye, Yankai Lin, Jiaju Du, Zhenghao Liu, Peng Li, Maosong Sun and Zhiyuan Liu ...	7170
<i>Is Graph Structure Necessary for Multi-hop Question Answering?</i>	
Nan Shao, Yiming Cui, Ting Liu, Shijin Wang and Guoping Hu	7187
<i>XL-WiC: A Multilingual Benchmark for Evaluating Semantic Contextualization</i>	
Alessandro Raganato, Tommaso Pasini, Jose Camacho-Collados and Mohammad Taher Pilehvar	7193
<i>Generatory or: "How we Went Beyond Word Sense Inventories and Learned to Gloss"</i>	
Michele Bevilacqua, Marco Maru and Roberto Navigli	7207
<i>Probing Pretrained Language Models for Lexical Semantics</i>	
Ivan Vulić, Edoardo Maria Ponti, Robert Litschko, Goran Glavaš and Anna Korhonen	7222
<i>Cross-lingual Spoken Language Understanding with Regularized Representation Alignment</i>	
Zihan Liu, Genta Indra Winata, Peng Xu, Zhaojiang Lin and Pascale Fung	7241
<i>SLURP: A Spoken Language Understanding Resource Package</i>	
Emanuele Bastianelli, Andrea Vanzo, Pawel Swietojanski and Verena Rieser	7252

<i>Neural Conversational QA: Learning to Reason vs Exploiting Patterns</i>	
Nikhil Verma, Abhishek Sharma, Dhiraj Madan, Danish Contractor, Harshit Kumar and Sachindra Joshi	7263
<i>Counterfactual Generator: A Weakly-Supervised Method for Named Entity Recognition</i>	
Xiangji Zeng, Yunliang Li, Yuchen Zhai and Yin Zhang	7270
<i>Understanding Procedural Text using Interactive Entity Networks</i>	
Jizhi Tang, Yansong Feng and Dongyan Zhao	7281
<i>A Rigorous Study on Named Entity Recognition: Can Fine-tuning Pretrained Model Lead to the Promised Land?</i>	
Hongyu Lin, Yaojie Lu, Jialong Tang, Xianpei Han, Le Sun, Zhicheng Wei and Nicholas Jing Yuan	7291
<i>DyERNIE: Dynamic Evolution of Riemannian Manifold Embeddings for Temporal Knowledge Graph Completion</i>	
Zhen Han, Peng Chen, Yunpu Ma and Volker Tresp	7301
<i>Embedding Words in Non-Vector Space with Unsupervised Graph Learning</i>	
Max Ryabinin, Sergei Popov, Liudmila Prokhorenkova and Elena Voita	7317
<i>Debiasing knowledge graph embeddings</i>	
Joseph Fisher, Arpit Mittal, Dave Palfrey and Christos Christodoulopoulos	7332
<i>Message Passing for Hyper-Relational Knowledge Graphs</i>	
Mikhail Galkin, Priyansh Trivedi, Gaurav Maheshwari, Ricardo Usbeck and Jens Lehmann ..	7346
<i>Relation-aware Graph Attention Networks with Relational Position Encodings for Emotion Recognition in Conversations</i>	
Taichi Ishiwatari, Yuki Yasuda, Taro Miyazaki and Jun Goto	7360
<i>BERT Knows Punta Cana is not just beautiful, it's gorgeous: Ranking Scalar Adjectives with Contextualised Representations</i>	
Aina Garí Soler and Marianna Apidianaki	7371
<i>Feature Adaptation of Pre-Trained Language Models across Languages and Domains with Robust Self-Training</i>	
Hai Ye, Qingyu Tan, Ruidan He, Juntao Li, Hwee Tou Ng and Lidong Bing	7386
<i>Textual Data Augmentation for Efficient Active Learning on Tiny Datasets</i>	
Husam Quteineh, Spyridon Samothrakis and Richard Sutcliffe	7400
<i>"I'd rather just go to bed": Understanding Indirect Answers</i>	
Annie Louis, Dan Roth and Filip Radlinski	7411
<i>PowerTransformer: Unsupervised Controllable Revision for Biased Language Correction</i>	
Xinyao Ma, Maarten Sap, Hannah Rashkin and Yejin Choi	7426
<i>MEGA RST Discourse Treebanks with Structure and Nuclearity from Scalable Distant Sentiment Supervision</i>	
Patrick Huber and Giuseppe Carenini	7442
<i>A Centering Approach for Discourse Structure-aware Coherence Modeling</i>	
Sungho Jeon and Michael Strube	7458

<i>Keeping Up Appearances: Computational Modeling of Face Acts in Persuasion Oriented Discussions</i> Ritam Dutt, Rishabh Joshi and Carolyn Rose	7473
<i>HABERTOR: An Efficient and Effective Deep Hatespeech Detector</i> Thanh Tran, Yifan Hu, Changwei Hu, Kevin Yen, Fei Tan, Kyumin Lee and Se Rim Park	7486
<i>An Empirical Study on Large-Scale Multi-Label Text Classification including Few and Zero-Shot Labels</i> Ilias Chalkidis, Manos Fergadiotis, Sotiris Kotitsas, Prodromos Malakasiotis, Nikolaos Aletras and Ion Androutsopoulos	7503
<i>Which *BERT? A Survey Organizing Contextualized Encoders</i> Patrick Xia, Shijie Wu and Benjamin Van Durme	7516
<i>Fact or Fiction: Verifying Scientific Claims</i> David Wadden, Shanchuan Lin, Kyle Lo, Lucy Lu Wang, Madeleine van Zuylen, Arman Cohan and Hannaneh Hajishirzi	7534
<i>Semantic Role Labeling as Syntactic Dependency Parsing</i> Tianze Shi, Igor Malioutov and Ozan Irsoy	7551
<i>PARADE: A New Dataset for Paraphrase Identification Requiring Computer Science Domain Knowledge</i> Yun He, Zhuoer Wang, Yin Zhang, Ruihong Huang and James Caverlee	7572
<i>Causal Inference of Script Knowledge</i> Noah Weber, Rachel Rudinger and Benjamin Van Durme	7583
<i>Towards Debiasing NLU Models from Unknown Biases</i> Prasetya Ajie Utama, Nafise Sadat Moosavi and Iryna Gurevych	7597
<i>On the Role of Supervision in Unsupervised Constituency Parsing</i> Haoyue Shi, Karen Livescu and Kevin Gimpel	7611
<i>Language Model Prior for Low-Resource Neural Machine Translation</i> Christos Baziotis, Barry Haddow and Alexandra Birch	7622
<i>Towards Detecting and Exploiting Disambiguation Biases in Neural Machine Translation</i> Denis Emelin, Ivan Titov and Rico Sennrich	7635
<i>MAD-X: An Adapter-based Framework for Multi-task Cross-lingual Transfer</i> Jonas Pfeiffer, Ivan Vulić, Iryna Gurevych and Sebastian Ruder	7654
<i>Translation Artifacts in Cross-lingual Transfer Learning</i> Mikel Artetxe, Gorka Labaka and Eneko Agirre	7674
<i>A Time-Aware Transformer based model for Suicide Ideation Detection on Social Media</i> Ramit Sawhney, Harshit Joshi, Saumya Gandhi and Rajiv Ratn Shah	7685
<i>Weakly Supervised Learning of Nuanced Frames for Analyzing Polarization in News Media</i> Shamik Roy and Dan Goldwasser	7698
<i>Where are the Facts? Searching for Fact-checked Information to Alleviate the Spread of Fake News</i> Nguyen Vo and Kyumin Lee	7717
<i>Fortifying Toxic Speech Detectors Against Disguised Toxicity</i> Xiaochuang Han and Yulia Tsvetkov	7732

<i>Explainable Automated Fact-Checking for Public Health Claims</i>	
Neema Kotonya and Francesca Toni	7740
<i>Interactive Fiction Game Playing as Multi-Paragraph Reading Comprehension with Reinforcement Learning</i>	
Xiaoxiao Guo, Mo Yu, Yupeng Gao, Chuang Gan, Murray Campbell and Shiyu Chang	7755
<i>DORB: Dynamically Optimizing Multiple Rewards with Bandits</i>	
Ramakanth Pasunuru, Han Guo and Mohit Bansal	7766
<i>Improving Detection and Categorization of Task-relevant Utterances through Integration of Discourse Structure and Ontological Knowledge</i>	
Sopan Khosla, Shikhar Vashishth, Jill Fain Lehman and Carolyn Rose	7781
<i>Hierarchical Evidence Set Modeling for Automated Fact Extraction and Verification</i>	
Shyam Subramanian and Kyumin Lee	7798
<i>Program Enhanced Fact Verification with Verbalization and Graph Attention Network</i>	
Xiaoyu Yang, Feng Nie, Yufei Feng, Quan Liu, Zhigang Chen and Xiaodan Zhu	7810
<i>Constrained Fact Verification for FEVER</i>	
Adithya Pratapa, Sai Muralidhar Jayanthi and Kavya Nerella	7826
<i>Entity Linking in 100 Languages</i>	
Jan A. Botha, Zifei Shan and Daniel Gillick	7833
<i>PatchBERT: Just-in-Time, Out-of-Vocabulary Patching</i>	
Sangwhan Moon and Naoaki Okazaki	7846
<i>On the importance of pre-training data volume for compact language models</i>	
Vincent Micheli, Martin d’Hoffschmidt and François Fleuret	7853
<i>BERT-of-Theseus: Compressing BERT by Progressive Module Replacing</i>	
Canwen Xu, Wangchunshu Zhou, Tao Ge, Furu Wei and Ming Zhou	7859
<i>Recall and Learn: Fine-tuning Deep Pretrained Language Models with Less Forgetting</i>	
Sanyuan Chen, Yutai Hou, Yiming Cui, Wanxiang Che, Ting Liu and Xiangzhan Yu	7870
<i>Exploring and Predicting Transferability across NLP Tasks</i>	
Tu Vu, Tong Wang, Tsendsuren Munkhdalai, Alessandro Sordoni, Adam Trischler, Andrew Mattarella-Micke, Subhransu Maji and Mohit Iyyer	7882
<i>To BERT or Not to BERT: Comparing Task-specific and Task-agnostic Semi-Supervised Approaches for Sequence Tagging</i>	
Kasturi Bhattacharjee, Miguel Ballesteros, Rishita Anubhai, Smaranda Muresan, Jie Ma, Faisal Ladhak and Yaser Al-Onaizan	7927
<i>Cold-start Active Learning through Self-Supervised Language Modeling</i>	
Michelle Yuan, Hsuan-Tien Lin and Jordan Boyd-Graber	7935
<i>Active Learning for BERT: An Empirical Study</i>	
Liat Ein-Dor, Alon Halfon, Ariel Gera, Eyal Shnarch, Lena Dankin, Leshem Choshen, Marina Danilevsky, Ranit Aharonov, Yoav Katz and Noam Slonim	7949
<i>Transformer Based Multi-Source Domain Adaptation</i>	
Dustin Wright and Isabelle Augenstein	7963

<i>Vector-Vector-Matrix Architecture: A Novel Hardware-Aware Framework for Low-Latency Inference in NLP Applications</i>	
Matthew Khoury, Rumén Dangovski, Longwu Ou, Preslav Nakov, Yichen Shen and Li Jing . . .	7975
<i>The importance of fillers for text representations of speech transcripts</i>	
Tanvi Dinkar, Pierre Colombo, Matthieu Labeau and Chloé Clavel	7985
<i>The role of context in neural pitch accent detection in English</i>	
Elizabeth Nielsen, Mark Steedman and Sharon Goldwater	7994
<i>VolTAGE: Volatility forecasting via Text-Audio fusion with Graph convolution networks for Earnings calls</i>	
Ramit Sawhney, Piyush Khanna, Arshiya Aggarwal, Taru Jain, Puneet Mathur and Rajiv Ratn Shah	8001
<i>Effectively pretraining a speech translation decoder with machine translation data</i>	
Ashkan Alinejad and Anoop Sarkar	8014
<i>A Preliminary Exploration of GANs for Keyphrase Generation</i>	
Avinash Swaminathan, Haimin Zhang, Debanjan Mahata, Rakesh Gosangi, Rajiv Ratn Shah and Amanda Stent	8021
<i>TESA: A Task in Entity Semantic Aggregation for Abstractive Summarization</i>	
Clément Jumel, Annie Louis and Jackie Chi Kit Cheung	8031
<i>MLSUM: The Multilingual Summarization Corpus</i>	
Thomas Scialom, Paul-Alexis Dray, Sylvain Lamprier, Benjamin Piwowarski and Jacopo Staiano	8051
<i>Multi-XScience: A Large-scale Dataset for Extreme Multi-document Summarization of Scientific Articles</i>	
Yao Lu, Yue Dong and Laurent Charlin	8068
<i>Intrinsic Evaluation of Summarization Datasets</i>	
Rishi Bommasani and Claire Cardie	8075
<i>Iterative Feature Mining for Constraint-Based Data Collection to Increase Data Diversity and Model Robustness</i>	
Stefan Larson, Anthony Zheng, Anish Mahendran, Rishi Tekriwal, Adrian Cheung, Eric Guldan, Kevin Leach and Jonathan K. Kummerfeld	8097
<i>Conversational Semantic Parsing for Dialog State Tracking</i>	
Jianpeng Cheng, Devang Agrawal, Héctor Martínez Alonso, Shruti Bhargava, Joris Driesen, Federico Flego, Dain Kaplan, Dimitri Katsaklis, Lin Li, Dhivya Piraviperumal, Jason D Williams, Hong Yu, Diarmuid Ó Séaghdha and Anders Johannsen	8107
<i>doc2dial: A Goal-Oriented Document-Grounded Dialogue Dataset</i>	
Song Feng, Hui Wan, Chulaka Gunasekara, Siva Patel, Sachindra Joshi and Luis Lastras	8118
<i>Interview: Large-scale Modeling of Media Dialog with Discourse Patterns and Knowledge Grounding</i>	
Bodhisattwa Prasad Majumder, Shuyang Li, Jianmo Ni and Julian McAuley	8129
<i>INSPIRED: Toward Sociable Recommendation Dialog Systems</i>	
Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyang Shi and Zhou Yu	8142

<i>Information Seeking in the Spirit of Learning: a Dataset for Conversational Curiosity</i>	
Pedro Rodriguez, Paul Crook, Seungwhan Moon and Zhiguang Wang	8153
<i>Queens are Powerful too: Mitigating Gender Bias in Dialogue Generation</i>	
Emily Dinan, Angela Fan, Adina Williams, Jack Urbanek, Douwe Kiela and Jason Weston ..	8173
<i>Discriminatively-Tuned Generative Classifiers for Robust Natural Language Inference</i>	
Xiaoan Ding, Tianyu Liu, Baobao Chang, Zhifang Sui and Kevin Gimpel	8189
<i>Collecting Entailment Data for Pretraining: New Protocols and Negative Results</i>	
Samuel R. Bowman, Jennimaria Palomaki, Livio Baldini Soares and Emily Pitler	8203
<i>The Curse of Performance Instability in Analysis Datasets: Consequences, Source, and Suggestions</i>	
Xiang Zhou, Yixin Nie, Hao Tan and Mohit Bansal	8215
<i>Universal Natural Language Processing with Limited Annotations: Try Few-shot Textual Entailment as a Start</i>	
Wenpeng Yin, Nazneen Fatema Rajani, Dragomir Radev, Richard Socher and Caiming Xiong	8229
<i>ConjNLI: Natural Language Inference Over Conjunctive Sentences</i>	
Swarnadeep Saha, Yixin Nie and Mohit Bansal	8240
<i>Data and Representation for Turkish Natural Language Inference</i>	
Emrah Budur, Rıza Özçelik, Tunga Gungor and Christopher Potts	8253
<i>Multitask Learning for Cross-Lingual Transfer of Broad-coverage Semantic Dependencies</i>	
Maryam Aminian, Mohammad Sadegh Rasooli and Mona Diab	8268
<i>Precise Task Formalization Matters in Winograd Schema Evaluations</i>	
Haokun Liu, William Huang, Dhara Mungra and Samuel R. Bowman	8275
<i>Gone At Last: Removing the Hypothesis-Only Bias in Natural Language Inference via Ensemble Adversarial Training</i>	
Joe Stacey, Pasquale Minervini, Haim Dubossarsky, Sebastian Riedel and Tim Rocktäschel ..	8281
<i>SynSetExpan: An Iterative Framework for Joint Entity Set Expansion and Synonym Discovery</i>	
Jiaming Shen, Wenda Qiu, Jingbo Shang, Michelle Vanni, Xiang Ren and Jiawei Han	8292
<i>Evaluating the Calibration of Knowledge Graph Embeddings for Trustworthy Link Prediction</i>	
Tara Safavi, Danai Koutra and Edgar Meij	8308
<i>Text Graph Transformer for Document Classification</i>	
Haopeng Zhang and Jiawei Zhang	8322
<i>CoDEx: A Comprehensive Knowledge Graph Completion Benchmark</i>	
Tara Safavi and Danai Koutra	8328
<i>META: Metadata-Empowered Weak Supervision for Text Classification</i>	
Dheeraj Mekala, Xinyang Zhang and Jingbo Shang	8351
<i>Towards More Accurate Uncertainty Estimation In Text Classification</i>	
Jianfeng He, Xuchao Zhang, Shuo Lei, Zhiqian Chen, Fanglan Chen, Abdulaziz Alhamadani, Bei Xiao and ChangTien Lu	8362
<i>Chapter Captor: Text Segmentation in Novels</i>	
Charuta Pethe, Allen Kim and Steve Skiena	8373

<i>Authorship Attribution for Neural Text Generation</i>	
Adaku Uchendu, Thai Le, Kai Shu and Dongwon Lee	8384
<i>NwQM: A neural quality assessment framework for Wikipedia</i>	
Bhanu Prakash Reddy Guda, Sasi Bhushan Seelaboyina, Soumya Sarkar and Animesh Mukherjee	8396
<i>Towards Modeling Revision Requirements in wikiHow Instructions</i>	
Irshad Bhat, Talita Anthonio and Michael Roth	8407
<i>Deep Attentive Learning for Stock Movement Prediction From Social Media Text and Company Correlations</i>	
Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa and Rajiv Ratn Shah	8415
<i>Natural Language Processing for Achieving Sustainable Development: the Case of Neural Labelling to Enhance Community Profiling</i>	
Costanza Conforti, Stephanie Hirmer, Dai Morgan, Marco Basaldella and Yau Ben Or	8427
<i>To Schedule or not to Schedule: Extracting Task Specific Temporal Entities and Associated Negation Constraints</i>	
Barun Patra, Chala Fufa, Pamela Bhattacharya and Charles Lee	8445
<i>Competence-Level Prediction and Resume-Job_Description Matching Using Context-Aware Transformer Models</i>	
Changmao Li, Elaine Fisher, Rebecca Thomas, Steve Pittard, Vicki Hertzberg and Jinho D. Choi	8456
<i>Grammatical Error Correction in Low Error Density Domains: A New Benchmark and Analyses</i>	
Simon Flachs, Ophélie Lacroix, Helen Yannakoudakis, Marek Rei and Anders Søgaard	8467
<i>Deconstructing word embedding algorithms</i>	
Kian Kenyon-Dean, Edward Newell and Jackie Chi Kit Cheung	8479
<i>Sequential Modelling of the Evolution of Word Representations for Semantic Change Detection</i>	
Adam Tsakalidis and Maria Liakata	8485
<i>Sparsity Makes Sense: Word Sense Disambiguation Using Sparse Contextualized Word Representations</i>	
Gábor Berend	8498
<i>Exploring Semantic Capacity of Terms</i>	
Jie Huang, Zilong Wang, Kevin Chang, Wen-mei Hwu and JinJun Xiong	8509
<i>Learning to Ignore: Long Document Coreference with Bounded Memory Neural Networks</i>	
Shubham Toshniwal, Sam Wiseman, Allyson Ettinger, Karen Livescu and Kevin Gimpel	8519
<i>Revealing the Myth of Higher-Order Inference in Coreference Resolution</i>	
Liyan Xu and Jinho D. Choi	8527
<i>Pre-training of Mention Representations in Coreference Models</i>	
Yuval Varkel and Amir Globerson	8534
<i>Learning Collaborative Agents with Rule Guidance for Knowledge Graph Reasoning</i>	
Deren Lei, Gangrong Jiang, Xiaotao Gu, Kexuan Sun, Yuning Mao and Xiang Ren	8541
<i>Exploring Contextualized Neural Language Models for Temporal Dependency Parsing</i>	
Hayley Ross, Jonathon Cai and Bonan Min	8548

<i>Systematic Comparison of Neural Architectures and Training Approaches for Open Information Extraction</i>	Patrick Hohenecker, Frank Mtumbuka, Vid Kocijan and Thomas Lukasiewicz	8554
<i>SeqMix: Augmenting Active Sequence Labeling via Sequence Mixup</i>	Rongzhi Zhang, Yue Yu and Chao Zhang	8566
<i>AxCell: Automatic Extraction of Results from Machine Learning Papers</i>	Marcin Kardas, Piotr Czapla, Pontus Stenetorp, Sebastian Ruder, Sebastian Riedel, Ross Taylor and Robert Stojnic	8580
<i>Knowledge-guided Open Attribute Value Extraction with Reinforcement Learning</i>	Ye Liu, Sheng Zhang, Rui Song, Suo Feng and Yanghua Xiao	8595
<i>DualTKB: A Dual Learning Bridge between Text and Knowledge Base</i>	Pierre Dognin, Igor Melnyk, Inkit Padhi, Cicero Nogueira dos Santos and Payel Das	8605
<i>Incremental Neural Coreference Resolution in Constant Memory</i>	Patrick Xia, João Sedoc and Benjamin Van Durme	8617
<i>Improving Low Compute Language Modeling with In-Domain Embedding Initialisation</i>	Charles Welch, Rada Mihalcea and Jonathan K. Kummerfeld	8625
<i>KGLM: Pretrained Knowledge-Grounded Language Model for Data-to-Text Generation</i>	Wenhu Chen, Yu Su, Xifeng Yan and William Yang Wang	8635
<i>Pointer: Constrained Text Generation via Insertion-based Generative Pre-training</i>	Yizhe Zhang, Guoyin Wang, Chunyuan Li, Zhe Gan, Chris Brockett and Bill Dolan	8649
<i>Unsupervised Text Style Transfer with Masked Language Models</i>	Eric Malmi, Aliaksei Severyn and Sascha Rothe	8671
<i>PALM: Pre-training an Autoencoding&autoregressive Language Model for Context-conditioned Generation</i>	Bin Bi, Chenliang Li, Chen Wu, Ming Yan, Wei Wang, Songfang Huang, Fei Huang and Luo Si	8681
<i>Gradient-guided Unsupervised Lexically Constrained Text Generation</i>	Lei Sha	8692
<i>TeaForN: Teacher-Forcing with N-grams</i>	Sebastian Goodman, Nan Ding and Radu Soricut	8704
<i>Experience Grounds Language</i>	Yonatan Bisk, Ari Holtzman, Jesse Thomason, Jacob Andreas, Yoshua Bengio, Joyce Chai, Mirella Lapata, Angeliki Lazaridou, Jonathan May, Aleksandr Nisnevich, Nicolas Pinto and Joseph Turian	8718
<i>Keep CALM and Explore: Language Models for Action Generation in Text-based Games</i>	Shunyu Yao, Rohan Rao, Matthew Hausknecht and Karthik Narasimhan	8736
<i>CapWAP: Captioning with a Purpose</i>	Adam Fisch, Kenton Lee, Ming-Wei Chang, Jonathan Clark and Regina Barzilay	8755
<i>What is More Likely to Happen Next? Video-and-Language Future Event Prediction</i>	Jie Lei, Licheng Yu, Tamara Berg and Mohit Bansal	8769

<i>X-LXMERT: Paint, Caption and Answer Questions with Multi-Modal Transformers</i>	
Jaemin Cho, jiasen lu, Dustin Schwenk, Hannaneh Hajishirzi and Aniruddha Kembhavi	8785
<i>Towards Understanding Sample Variance in Visually Grounded Language Generation: Evaluations and Observations</i>	
Wanrong Zhu, Xin Wang, Pradyumna Narayana, Kazoo Sone, Sugato Basu and William Yang Wang	8806
<i>Beyond Instructional Videos: Probing for More Diverse Visual-Textual Grounding on YouTube</i>	
Jack Hessel, Zhenhai Zhu, Bo Pang and Radu Soricut	8812
<i>Hierarchical Graph Network for Multi-hop Question Answering</i>	
Yuwei Fang, Siqi Sun, Zhe Gan, Rohit Pillai, Shuohang Wang and Jingjing Liu	8823
<i>A Simple Yet Strong Pipeline for HotpotQA</i>	
Dirk Groeneveld, Tushar Khot, and Ashish Sabharwal	8839
<i>Is Multihop QA in DiRe Condition? Measuring and Reducing Disconnected Reasoning</i>	
Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot and Ashish Sabharwal	8846
<i>Unsupervised Question Decomposition for Question Answering</i>	
Ethan Perez, Patrick Lewis, Wen-tau Yih, Kyunghyun Cho and Douwe Kiela	8864
<i>SRLGRN: Semantic Role Labeling Graph Reasoning Network</i>	
Chen Zheng and Parisa Kordjamshidi	8881
<i>CancerEmo: A Dataset for Fine-Grained Emotion Detection</i>	
Tiberiu Sosea and Cornelia Caragea	8892
<i>Exploring the Role of Argument Structure in Online Debate Persuasion</i>	
Jialu Li, Esin Durmus and Claire Cardie	8905
<i>Zero-Shot Stance Detection: A Dataset and Model using Generalized Topic Representations</i>	
Emily Allaway and Kathleen McKeown	8913
<i>Sentiment Analysis of Tweets using Heterogeneous Multi-layer Network Representation and Embedding</i>	
Loitongbam Gyanendro Singh, Anasua Mitra and Sanasam Ranbir Singh	8932
<i>Introducing Syntactic Structures into Target Opinion Word Extraction with Deep Learning</i>	
Amir Pouran Ben Veyseh, Nasim Nouri, Franck Dernoncourt, Dejing Dou and Thien Huu Nguyen	8947
<i>Can Emojis Convey Human Emotions? A Study to Understand the Association between Emojis and Emotions</i>	
Abu Awal Md Shoeb and Gerard de Melo	8957
<i>MIME: MIMicking Emotions for Empathetic Response Generation</i>	
Navonil Majumder, Pengfei Hong, Shanshan Peng, Jiankun Lu, Deepanway Ghosal, Alexander Gelbukh, Rada Mihalcea and Soujanya Poria	8968
<i>Exploiting Structured Knowledge in Text via Graph-Guided Representation Learning</i>	
Tao Shen, Yi Mao, Pengcheng He, Guodong Long, Adam Trischler and Weizhu Chen	8980
<i>Named Entity Recognition Only from Word Embeddings</i>	
Ying Luo, Hai Zhao and Junlang Zhan	8995

<i>Weakly-Supervised Text Classification Using Label Names Only</i>	
Yu Meng, Yunyi Zhang, Jiaxin Huang, Chenyan Xiong, Heng Ji, Chao Zhang and Jiawei Han	9006
<i>Neural Topic Modeling with Cycle-Consistent Adversarial Training</i>	
Xuemeng Hu, Rui Wang, Deyu Zhou and Yuxuan Xiong	9018
<i>Data Boost: Text Data Augmentation Through Reinforcement Learning Guided Conditional Generation</i>	
Ruibo Liu, Guangxuan Xu, Chenyan Jia, Weicheng Ma, Lili Wang and Soroush Vosoughi	9031
<i>A State-independent and Time-evolving Network with Applications to Early Rumor Detection</i>	
Rui Xia, Kaizhou Xuan and Jianfei Yu	9042
<i>PyMT5: multi-mode translation of natural language and Python code with transformers</i>	
Colin Clement, Dawn Drain, Jonathan Timcheck, Alexey Svyatkovskiy and Neel Sundaresan	9052
<i>PathQG: Neural Question Generation from Facts</i>	
Siyuan Wang, Zhongyu Wei, Zhihao Fan, Zengfeng Huang, Weijian Sun, Qi ZHANG and Xuanjing Huang	9066
<i>What time is it? Temporal Analysis of Novels</i>	
Allen Kim, Charuta Pethe and Steve Skiena	9076
<i>COGS: A Compositional Generalization Challenge Based on Semantic Interpretation</i>	
Najoung Kim and Tal Linzen	9087
<i>An Analysis of Natural Language Inference Benchmarks through the Lens of Negation</i>	
Md Mosharaf Hossain, Venelin Kovatchev, Pranoy Dutta, Tiffany Kao, Elizabeth Wei and Eduardo Blanco	9106
<i>On the Sentence Embeddings from BERT for Semantic Textual Similarity</i>	
Bohan Li, Hao Zhou, Junxian He, Mingxuan Wang, Yiming Yang and Lei Li	9119
<i>What Can We Learn from Collective Human Opinions on Natural Language Inference Data?</i>	
Yixin Nie, Xiang Zhou and Mohit Bansal	9131
<i>Improving Text Generation with Student-Forcing Optimal Transport</i>	
Jianqiao Li, Chunyuan Li, Guoyin Wang, Hao Fu, Yuhchen Lin, Liqun Chen, Yizhe Zhang, Chenyang Tao, Ruiyi Zhang, Wenlin Wang, Dinghan Shen, Qian Yang and Lawrence Carin	9144
<i>UNION: An Unreferenced Metric for Evaluating Open-ended Story Generation</i>	
Jian Guan and Minlie Huang	9157
<i>F²-Softmax: Diversifying Neural Text Generation via Frequency Factorized Softmax</i>	
Byung-Ju Choi, Jimin Hong, David Park and Sang Wan Lee	9167
<i>Partially-Aligned Data-to-Text Generation with Distant Supervision</i>	
Zihao Fu, Bei Shi, Wai Lam, Lidong Bing and Zhiyuan Liu	9183
<i>Like hiking? You probably enjoy nature: Persona-grounded Dialog with Commonsense Expansions</i>	
Bodhisattwa Prasad Majumder, Harsh Jhamtani, Taylor Berg-Kirkpatrick and Julian McAuley	9194
<i>A Probabilistic End-To-End Task-Oriented Dialog Model with Latent Belief States towards Semi-Supervised Learning</i>	
Yichi Zhang, Zhijian Ou, Min Hu and Junlan Feng	9207

<i>The World is not Binary: Learning to Rank with Grayscale Data for Dialogue Response Selection</i>	
Zibo Lin, Deng Cai, Yan Wang, Xiaojiang Liu, Haitao Zheng and Shuming Shi	9220
<i>GRADE: Automatic Graph-Enhanced Coherence Metric for Evaluating Open-Domain Dialogue Systems</i>	
Lishan Huang, Zheng Ye, Jinghui Qin, Liang Lin and Xiaodan Liang	9230
<i>MedDialog: A Large-scale Medical Dialogue Dataset</i>	
Guangtao Zeng, Wenmian Yang, Zeqian Ju, Yue Yang, Sicheng Wang, Ruisi Zhang, Meng Zhou, Jiaqi Zeng, Xiangyu Dong, Ruoyu Zhang, Hongchao Fang, Penghui Zhu, Shu Chen and Pengtao Xie	9241
<i>An information theoretic view on selecting linguistic probes</i>	
Zining Zhu and Frank Rudzicz	9251
<i>With Little Power Comes Great Responsibility</i>	
Dallas Card, Peter Henderson, Urvashi Khandelwal, Robin Jia, Kyle Mahowald and Dan Jurafsky	9263
<i>Dataset Cartography: Mapping and Diagnosing Datasets with Training Dynamics</i>	
Swabha Swayamdipta, Roy Schwartz, Nicholas Lourie, Yizhong Wang, Hannaneh Hajishirzi, Noah A. Smith and Yejin Choi	9275
<i>Evaluating and Characterizing Human Rationales</i>	
Samuel Carton, Anirudh Rathore and Chenhao Tan	9294
<i>On Extractive and Abstractive Neural Document Summarization with Transformer Language Models</i>	
Jonathan Pilault, Raymond Li, Sandeep Subramanian and Chris Pal	9308
<i>Multi-Fact Correction in Abstractive Text Summarization</i>	
Yue Dong, Shuohang Wang, Zhe Gan, Yu Cheng, Jackie Chi Kit Cheung and Jingjing Liu . . .	9320
<i>Evaluating the Factual Consistency of Abstractive Text Summarization</i>	
Wojciech Kryscinski, Bryan McCann, Caiming Xiong and Richard Socher	9332
<i>Re-evaluating Evaluation in Text Summarization</i>	
Manik Bhandari, Pranav Narayan Gour, Atabak Ashfaq, Pengfei Liu and Graham Neubig . . .	9347
<i>VMSMO: Learning to Generate Multimodal Summary for Video-based News Articles</i>	
Mingzhe Li, Xiuying Chen, Shen Gao, Zhangming Chan, Dongyan Zhao and Rui Yan	9360