

# **58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)**

Online  
5 – 10 July 2020

Volume 1 of 17

ISBN: 978-1-7138-1371-2

**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. (2020)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Learning to Understand Child-directed and Adult-directed Speech</i> Lieke Gelderloos, Grzegorz Chrupala and Afra Alishahi .....	1
<i>Predicting Depression in Screening Interviews from Latent Categorization of Interview Prompts</i> Alex Rinaldi, Jean Fox Tree and Snigdha Chaturvedi .....	7
<i>Coach: A Coarse-to-Fine Approach for Cross-domain Slot Filling</i> Zihan Liu, Genta Indra Winata, Peng Xu and Pascale Fung .....	19
<i>Designing Precise and Robust Dialogue Response Evaluators</i> Tianyu Zhao, Divesh Lala and Tatsuya Kawahara .....	26
<i>Dialogue State Tracking with Explicit Slot Connection Modeling</i> Yawen Ouyang, Moxin Chen, Xinyu Dai, Yinggong Zhao, Shujian Huang and Jiajun CHEN ...	34
<i>Generating Informative Conversational Response using Recurrent Knowledge-Interaction and Knowledge-Copy</i> Xiexiong Lin, Weiyu Jian, Jianshan He, Taifeng Wang and Wei Chu .....	41
<i>Guiding Variational Response Generator to Exploit Persona</i> Bowen Wu, MengYuan Li, Zongsheng Wang, Yifu Chen, Derek F. Wong, qihang feng, Junhong Huang and Baoxun Wang .....	53
<i>Large Scale Multi-Actor Generative Dialog Modeling</i> Alex Boyd, Raul Puri, Mohammad Shoeybi, Mostofa Patwary and Bryan Catanzaro .....	66
<i>PLATO: Pre-trained Dialogue Generation Model with Discrete Latent Variable</i> Siqi Bao, Huang He, Fan Wang, Hua Wu and Haifeng Wang .....	85
<i>Slot-consistent NLG for Task-oriented Dialogue Systems with Iterative Rectification Network</i> Yangming Li, Kaisheng Yao, Libo Qin, Wanxiang Che, Xiaolong Li and Ting Liu .....	97
<i>Span-CONVERT: Few-shot Span Extraction for Dialog with Pretrained Conversational Representations</i> Samuel Coope, Tyler Farghly, Daniela Gerz, Ivan Vulić and Matthew Henderson .....	107
<i>Zero-Shot Transfer Learning with Synthesized Data for Multi-Domain Dialogue State Tracking</i> Giovanni Campagna, Agata Foryciarz, Mehrad Moradshahi and Monica Lam .....	122
<i>A Complete Shift-Reduce Chinese Discourse Parser with Robust Dynamic Oracle</i> Shyh-Shiun Hung, Hen-Hsen Huang and Hsin-Hsi Chen .....	133
<i>TransS-Driven Joint Learning Architecture for Implicit Discourse Relation Recognition</i> Ruifang He, Jian Wang, Fengyu Guo and Yugui Han .....	139
<i>A Study of Non-autoregressive Model for Sequence Generation</i> Yi Ren, Jinglin Liu, Xu Tan, Zhou Zhao, sheng zhao and Tie-Yan Liu .....	149
<i>Cross-modal Language Generation using Pivot Stabilization for Web-scale Language Coverage</i> Ashish V. Thapliyal and Radu Soricut .....	160
<i>Fact-based Text Editing</i> Hayate Iso, Chao Qiao and Hang Li .....	171

<i>Few-Shot NLG with Pre-Trained Language Model</i>	
Zhiyu Chen, Harini Eavani, Wenhui Chen, Yinyin Liu and William Yang Wang	183
<i>Fluent Response Generation for Conversational Question Answering</i>	
Ashutosh Baheti, Alan Ritter and Kevin Small	191
<i>Generating Diverse and Consistent QA pairs from Contexts with Information-Maximizing Hierarchical Conditional VAEs</i>	
Dong Bok Lee, Seanie Lee, Woo Tae Jeong, Donghwan Kim and Sung Ju Hwang	208
<i>Learning to Ask More: Semi-Autoregressive Sequential Question Generation under Dual-Graph Interaction</i>	
Zi Chai and Xiaojun Wan	225
<i>Neural Syntactic Preordering for Controlled Paraphrase Generation</i>	
Tanya Goyal and Greg Durrett	238
<i>Pre-train and Plug-in: Flexible Conditional Text Generation with Variational Auto-Encoders</i>	
Yu Duan, Canwen Xu, Jiaxin Pei, Jialong Han and Chenliang Li	253
<i>Probabilistically Masked Language Model Capable of Autoregressive Generation in Arbitrary Word Order</i>	
Yi Liao, Xin Jiang and Qun Liu	263
<i>Reverse Engineering Configurations of Neural Text Generation Models</i>	
Yi Tay, Dara Bahri, Che Zheng, Clifford Brunk, Donald Metzler and Andrew Tomkins	275
<i>Review-based Question Generation with Adaptive Instance Transfer and Augmentation</i>	
Qian Yu, Lidong Bing, Qiong Zhang, Wai Lam and Luo Si	280
<i>TAG : Type Auxiliary Guiding for Code Comment Generation</i>	
Ruichu Cai, Zhihao Liang, Boyan Xu, zijian li, Yuxing Hao and Yao Chen	291
<i>Unsupervised Paraphrasing by Simulated Annealing</i>	
Xianggen Liu, Lili Mou, Fandong Meng, Hao Zhou, Jie Zhou and Sen Song	302
<i>A Joint Model for Document Segmentation and Segment Labeling</i>	
Joe Barrow, Rajiv Jain, Vlad Morariu, Varun Manjunatha, Douglas Oard and Philip Resnik	313
<i>Contextualized Weak Supervision for Text Classification</i>	
Dheeraj Mekala and Jingbo Shang	323
<i>Every Document Owns Its Structure: Inductive Text Classification via Graph Neural Networks</i>	
Yufeng Zhang, Xueli Yu, Zeyu Cui, Shu Wu, Zhongzhen Wen and Liang Wang	334
<i>Neural Topic Modeling with Bidirectional Adversarial Training</i>	
Rui Wang, Xuemeng Hu, Deyu Zhou, Yulan He, Yuxuan Xiong, Chenchen Ye and Haiyang Xu	340
<i>Text Classification with Negative Supervision</i>	
Sora Ohashi, Junya Takayama, Tomoyuki Kajiwara, Chenhui Chu and Yuki Arase	351
<i>Content Word Aware Neural Machine Translation</i>	
Kehai Chen, Rui Wang, Masao Utiyama and Eiichiro Sumita	358
<i>Evaluating Explanation Methods for Neural Machine Translation</i>	
Jierui Li, Lemao Liu, Huayang Li, Guanlin Li, Guoping Huang and Shuming Shi	365

<i>Jointly Masked Sequence-to-Sequence Model for Non-Autoregressive Neural Machine Translation</i> Junliang Guo, Linli Xu and Enhong Chen .....	376
<i>Learning Source Phrase Representations for Neural Machine Translation</i> Hongfei Xu, Josef van Genabith, Deyi Xiong, Qiuhui Liu and Jingyi Zhang .....	386
<i>Lipschitz Constrained Parameter Initialization for Deep Transformers</i> Hongfei Xu, Qiuhui Liu, Josef van Genabith, Deyi Xiong and Jingyi Zhang .....	397
<i>Location Attention for Extrapolation to Longer Sequences</i> Yann Dubois, Gautier Dagan, Dieuwke Hupkes and Elia Bruni .....	403
<i>Multiscale Collaborative Deep Models for Neural Machine Translation</i> Xiangpeng Wei, Heng Yu, Yue Hu, Yue Zhang, Rongxiang Weng and Weihua Luo .....	414
<i>Norm-Based Curriculum Learning for Neural Machine Translation</i> Xuebo Liu, Houtim Lai, Derek F. Wong and Lidia S. Chao .....	427
<i>Opportunistic Decoding with Timely Correction for Simultaneous Translation</i> Renjie Zheng, Mingbo Ma, Baigong Zheng, Kaibo Liu and Liang Huang .....	437
<i>A Formal Hierarchy of RNN Architectures</i> William Merrill, Gail Weiss, Yoav Goldberg, Roy Schwartz, Noah A. Smith and Eran Yahav ..	443
<i>A Three-Parameter Rank-Frequency Relation in Natural Languages</i> Chenchen Ding, Masao Utiyama and Eiichiro Sumita .....	460
<i>Dice Loss for Data-imbalanced NLP Tasks</i> Xiaoya Li, Xiaofei Sun, Yuxian Meng, Junjun Liang, Fei Wu and Jiwei Li .....	465
<i>Emergence of Syntax Needs Minimal Supervision</i> Raphaël Bailly and Kata Gábor .....	477
<i>Language Models as an Alternative Evaluator of Word Order Hypotheses: A Case Study in Japanese</i> Tatsuki Kuribayashi, Takumi Ito, Jun Suzuki and Kentaro Inui .....	488
<i>GCAN: Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media</i> Yi-Ju Lu and Cheng-Te Li .....	505
<i>Integrating Semantic and Structural Information with Graph Convolutional Network for Controversy Detection</i> Lei Zhong, Juan Cao, Qiang Sheng, Junbo Guo and Ziang Wang .....	515
<i>Predicting the Topical Stance and Political Leaning of Media using Tweets</i> Peter Stefanov, Kareem Darwish, Atanas Atanasov and Preslav Nakov .....	527
<i>Simple, Interpretable and Stable Method for Detecting Words with Usage Change across Corpora</i> Hila Gonen, Ganesh Jawahar, Djamé Seddah and Yoav Goldberg .....	538
<i>CDL: Curriculum Dual Learning for Emotion-Controllable Response Generation</i> Lei Shen and Yang Feng .....	556
<i>Efficient Dialogue State Tracking by Selectively Overwriting Memory</i> Sungdong Kim, Sohee Yang, Gyuwan Kim and Sang-Woo Lee .....	567

<i>End-to-End Neural Pipeline for Goal-Oriented Dialogue Systems using GPT-2</i> Donghoon Ham, Jeong-Gwan Lee, Youngsoo Jang and Kee-Eung Kim .....	583
<i>Evaluating Dialogue Generation Systems via Response Selection</i> Shiki Sato, Reina Akama, Hiroki Ouchi, Jun Suzuki and Kentaro Inui .....	593
<i>Gated Convolutional Bidirectional Attention-based Model for Off-topic Spoken Response Detection</i> Yefei Zha, Ruobing Li and Hui Lin .....	600
<i>Learning Low-Resource End-To-End Goal-Oriented Dialog for Fast and Reliable System Deployment</i> Yinpei Dai, Hangyu Li, Chengguang Tang, Yongbin Li, Jian Sun and Xiaodan Zhu .....	609
<i>Learning to Tag OOV Tokens by Integrating Contextual Representation and Background Knowledge</i> Keqing He, Yuanmeng Yan and Weiran XU .....	619
<i>Multi-Agent Task-Oriented Dialog Policy Learning with Role-Aware Reward Decomposition</i> Ryuichi Takanobu, Runze Liang and Minlie Huang .....	625
<i>Paraphrase Augmented Task-Oriented Dialog Generation</i> Silin Gao, Yichi Zhang, Zhijian Ou and Zhou Yu .....	639
<i>Response-Anticipated Memory for On-Demand Knowledge Integration in Response Generation</i> Zhiliang Tian, Wei Bi, Dongkyu Lee, Lanqing Xue, YIPING SONG, Xiaojiang Liu and Nevin L. Zhang .....	650
<i>Semi-Supervised Dialogue Policy Learning via Stochastic Reward Estimation</i> Xinting Huang, Jianzhong Qi, Yu Sun and Rui Zhang .....	660
<i>Towards Unsupervised Language Understanding and Generation by Joint Dual Learning</i> Shang-Yu Su, Chao-Wei Huang and Yun-Nung Chen .....	671
<i>USR: An Unsupervised and Reference Free Evaluation Metric for Dialog Generation</i> Shikib Mehri and Maxine Eskenazi .....	681
<i>Explicit Semantic Decomposition for Definition Generation</i> Jiahuan Li, Yu Bao, Shujian Huang, Xinyu Dai and Jiajun CHEN .....	708
<i>Improved Natural Language Generation via Loss Truncation</i> Daniel Kang and Tatsunori Hashimoto .....	718
<i>Line Graph Enhanced AMR-to-Text Generation with Mix-Order Graph Attention Networks</i> Yanbin Zhao, Lu Chen, Zhi Chen, Ruisheng Cao, Su Zhu and Kai Yu .....	732
<i>Rigid Formats Controlled Text Generation</i> Piji Li, Haisong Zhang, Xiaojiang Liu and Shuming Shi .....	742
<i>Syn-QG: Syntactic and Shallow Semantic Rules for Question Generation</i> Kaustubh Dhole and Christopher D. Manning .....	752
<i>An Online Semantic-enhanced Dirichlet Model for Short Text Stream Clustering</i> Jay Kumar, Junming Shao, Salah Uddin and Wazir Ali .....	766
<i>Generative Semantic Hashing Enhanced via Boltzmann Machines</i> Lin Zheng, Qinliang Su, Dinghan Shen and Changyou Chen .....	777

<i>Interactive Construction of User-Centric Dictionary for Text Analytics</i>	
Ryosuke Kohita, Issei Yoshida, Hiroshi Kanayama and Tetsuya Nasukawa .....	789
<i>Tree-Structured Neural Topic Model</i>	
Masaru Isonuma, Junichiro Mori, Danushka Bollegala and Ichiro Sakata .....	800
<i>Unsupervised FAQ Retrieval with Question Generation and BERT</i>	
Yosi Mass, Boaz Carmeli, Haggai Roitman and David Konopnicki .....	807
<i>"The Boating Store Had Its Best Sail Ever": Pronunciation-attentive Contextualized Pun Recognition</i>	
Yichao Zhou, Jyun-Yu Jiang, Jieyu Zhao, Kai-Wei Chang and Wei Wang .....	813
<i>Fast and Accurate Deep Bidirectional Language Representations for Unsupervised Learning</i>	
Joongbo Shin, Yoonhyung Lee, Seunghyun Yoon and Kyomin Jung .....	823
<i>Fine-grained Interest Matching for Neural News Recommendation</i>	
Heyuan Wang, Fangzhao Wu, Zheng Liu and Xing Xie .....	836
<i>Interpretable Operational Risk Classification with Semi-Supervised Variational Autoencoder</i>	
Fan Zhou, Shengming Zhang and Yi Yang .....	846
<i>Interpreting Twitter User Geolocation</i>	
Ting Zhong, Tianliang Wang, Fan Zhou, Goce Trajcevski, Kunpeng Zhang and Yi Yang .....	853
<i>Modeling Code-Switch Languages Using Bilingual Parallel Corpus</i>	
Grandee Lee and Haizhou Li .....	860
<i>SpellGCN: Incorporating Phonological and Visual Similarities into Language Models for Chinese Spelling Check</i>	
Xingyi Cheng, Weidi Xu, Kunlong Chen, Shaohua Jiang, Feng Wang, Taifeng Wang, Wei Chu and Yuan Qi .....	871
<i>Spelling Error Correction with Soft-Masked BERT</i>	
Shaohua Zhang, Haoran Huang, Jicong Liu and Hang Li .....	882
<i>A Frame-based Sentence Representation for Machine Reading Comprehension</i>	
Shaoru Guo, Ru Li, Hongye Tan, Xiaoli Li, Yong Guan, Hongyan Zhao and Yueping Zhang ..	891
<i>A Methodology for Creating Question Answering Corpora Using Inverse Data Annotation</i>	
Jan Deriu, Katsiaryna Mlynchyk, Philippe Schläpfer, Alvaro Rodrigo, Dirk von Grünigen, Nicolas Kaiser, Kurt Stockinger, Eneko Agirre and Mark Cieliebak .....	897
<i>Contextualized Sparse Representations for Real-Time Open-Domain Question Answering</i>	
Jinhyuk Lee, Minjoon Seo, Hannaneh Hajishirzi and Jaewoo Kang .....	912
<i>Dynamic Sampling Strategies for Multi-Task Reading Comprehension</i>	
Ananth Gottumukkala, Dheeru Dua, Sameer Singh and Matt Gardner .....	920
<i>Enhancing Answer Boundary Detection for Multilingual Machine Reading Comprehension</i>	
Fei Yuan, Linjun Shou, Xuanyu Bai, Ming Gong, Yaobo Liang, Nan Duan, Yan Fu and Daxin Jiang	925
<i>Explicit Memory Tracker with Coarse-to-Fine Reasoning for Conversational Machine Reading</i>	
Yifan Gao, Chien-Sheng Wu, Shafiq Joty, Caiming Xiong, Richard Socher, Irwin King, Michael Lyu and Steven C.H. Hoi .....	935

<i>Injecting Numerical Reasoning Skills into Language Models</i>	
Mor Geva, Ankit Gupta and Jonathan Berant . . . . .	946
<i>Learning to Identify Follow-Up Questions in Conversational Question Answering</i>	
Souvik Kundu, Qian Lin and Hwee Tou Ng . . . . .	959
<i>Query Graph Generation for Answering Multi-hop Complex Questions from Knowledge Bases</i>	
Yunshi Lan and Jing Jiang . . . . .	969
<i>A Diverse Corpus for Evaluating and Developing English Math Word Problem Solvers</i>	
Shen-yun Miao, Chao-Chun Liang and Keh-Yih Su . . . . .	975
<i>Improving Image Captioning Evaluation by Considering Inter References Variance</i>	
Yanzhi Yi, Hangyu Deng and Jinglu Hu . . . . .	985
<i>Revisiting the Context Window for Cross-lingual Word Embeddings</i>	
Ryokan Ri and Yoshimasa Tsuruoka . . . . .	995
<i>Moving Down the Long Tail of Word Sense Disambiguation with Gloss Informed Bi-encoders</i>	
Terra Blevins and Luke Zettlemoyer . . . . .	1006
<i>Code-Switching Patterns Can Be an Effective Route to Improve Performance of Downstream NLP Applications: A Case Study of Humour, Sarcasm and Hate Speech Detection</i>	
Srijan Bansal, Vishal Garimella, Ayush Suhane, Jasabanta Patro and Animesh Mukherjee . . . . .	1018
<i>DTCA: Decision Tree-based Co-Attention Networks for Explainable Claim Verification</i>	
Lianwei Wu, Yuan Rao, yongqiang zhao, Hao Liang and Ambreen Nazir . . . . .	1024
<i>Towards Conversational Recommendation over Multi-Type Dialogs</i>	
Zeming Liu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che and Ting Liu . . . . .	1036
<i>Unknown Intent Detection Using Gaussian Mixture Model with an Application to Zero-shot Intent Classification</i>	
Guangfeng Yan, Lu Fan, Qimai Li, Han Liu, Xiaotong Zhang, Xiao-Ming Wu and Albert Y.S. Lam	1050
<i>Expertise Style Transfer: A New Task Towards Better Communication between Experts and Laymen</i>	
Yixin Cao, Ruihao Shui, Liangming Pan, Min-Yen Kan, Zhiyuan Liu and Tat-Seng Chua . . . . .	1061
<i>Towards Faithful Neural Table-to-Text Generation with Content-Matching Constraints</i>	
Zhenyi Wang, Xiaoyang Wang, Bang An, Dong Yu and Changyou Chen . . . . .	1072
<i>Dynamic Memory Induction Networks for Few-Shot Text Classification</i>	
Ruiying Geng, Binhua Li, Yongbin Li, Jian Sun and Xiaodan Zhu . . . . .	1087
<i>Exclusive Hierarchical Decoding for Deep Keyphrase Generation</i>	
Wang Chen, Hou Pong Chan, Piji Li and Irwin King . . . . .	1095
<i>Hierarchy-Aware Global Model for Hierarchical Text Classification</i>	
Jie Zhou, Chunping Ma, Dingkun Long, Guangwei Xu, Ning Ding, Haoyu Zhang, Pengjun Xie and Gongshen Liu . . . . .	1106
<i>Keyphrase Generation for Scientific Document Retrieval</i>	
Florian Boudin, Ygor Gallina and Akiko Aizawa . . . . .	1118



<i>A Graph Auto-encoder Model of Derivational Morphology</i>	
Valentin Hofmann, Hinrich Schütze and Janet Pierrehumbert .....	1127
<i>Building a User-Generated Content North-African Arabizi Treebank: Tackling Hell</i>	
Djamé Seddah, Farah Essaidi, Amal Fethi, Matthieu Futral, Benjamin Muller, Pedro Javier Ortiz Suárez, Benoît Sagot and Abhishek Srivastava .....	1139
<i>Crawling and Preprocessing Mailing Lists At Scale for Dialog Analysis</i>	
Janek Bevendorff, Khalid Al Khatib, Martin Potthast and Benno Stein .....	1151
<i>Fine-Grained Analysis of Cross-Linguistic Syntactic Divergences</i>	
Dmitry Nikolaev, Ofir Arviv, Taelin Karidi, Neta Kenneth, Veronika Mitnik, Lilja Maria Saeboe and Omri Abend .....	1159
<i>Generating Counter Narratives against Online Hate Speech: Data and Strategies</i>	
Serra Sinem Tekiroğlu, Yi-Ling Chung and Marco Guerini .....	1177
<i>KLEJ: Comprehensive Benchmark for Polish Language Understanding</i>	
Piotr Rybak, Robert Mroczkowski, Janusz Tracz and Ireneusz Gawlik .....	1191
<i>Learning and Evaluating Emotion Lexicons for 91 Languages</i>	
Sven Buechel, Susanna Rücker and Udo Hahn .....	1202
<i>Multi-Hypothesis Machine Translation Evaluation</i>	
Marina Fomicheva, Lucia Specia and Francisco Guzmán .....	1218
<i>Multimodal Quality Estimation for Machine Translation</i>	
Shu Okabe, Frédéric Blain and Lucia Specia .....	1233
<i>PuzzLing Machines: A Challenge on Learning From Small Data</i>	
Gözde Gül Şahin, Yova Kementchedjieva, Phillip Rust and Iryna Gurevych .....	1241
<i>The SOFC-Exp Corpus and Neural Approaches to Information Extraction in the Materials Science Domain</i>	
Annemarie Friedrich, Heike Adel, Federico Tomazic, Johannes Hingerl, Renou Benteau, Anika Maruszczyk and Lukas Lange .....	1255
<i>The TechQA Dataset</i>	
Vittorio Castelli, Rishav Chakravarti, Saswati Dana, Anthony Ferritto, Radu Florian, Martin Franz, Dinesh Garg, Dinesh Khandelwal, Scott McCarley, Michael McCawley, Mohamed Nasr, Lin Pan, Cezar Pendus, John Pitrelli, Saurabh Pujar, Salim Roukos, Andrzej Sakrajda, Avi Sil, Rosario Uceda-Sosa, Todd Ward and Rong Zhang .....	1269
<i>iSarcasm: A Dataset of Intended Sarcasm</i>	
Silviu Oprea and Walid Magdy .....	1279
<i>AMR Parsing via Graph-Sequence Iterative Inference</i>	
Deng Cai and Wai Lam .....	1290
<i>A Large-Scale Multi-Document Summarization Dataset from the Wikipedia Current Events Portal</i>	
Demian Gholipour Ghalandari, Chris Hokamp, Nghia The Pham, John Glover and Georgiana Ifrim	1302
<i>Attend, Translate and Summarize: An Efficient Method for Neural Cross-Lingual Summarization</i>	
Junnan Zhu, Yu Zhou, Jiajun Zhang and Chengqing Zong .....	1309

<i>Examining the State-of-the-Art in News Timeline Summarization</i> Demian Gholipour Ghalandari and Georgiana Ifrim .....	1322
<i>Improving Truthfulness of Headline Generation</i> Kazuki Matsumaru, Sho Takase and Naoaki Okazaki .....	1335
<i>SUPERT: Towards New Frontiers in Unsupervised Evaluation Metrics for Multi-Document Summarization</i> Yang Gao, Wei Zhao and Steffen Eger .....	1347
<i>Self-Attention Guided Copy Mechanism for Abstractive Summarization</i> Song Xu, Haoran Li, Peng Yuan, Youzheng Wu, Xiaodong He and Bowen Zhou .....	1355
<i>Beyond User Self-Reported Likert Scale Ratings: A Comparison Model for Automatic Dialog Evaluation</i> Weixin Liang, James Zou and Zhou Yu .....	1363
<i>Conversational Word Embedding for Retrieval-Based Dialog System</i> Wentao Ma, Yiming Cui, Ting Liu, Dong Wang, Shijin Wang and Guoping Hu .....	1375
<i>Few-shot Slot Tagging with Collapsed Dependency Transfer and Label-enhanced Task-adaptive Projection Network</i> Yutai Hou, Wanxiang Che, Yongkui Lai, Zhihan Zhou, Yijia Liu, Han Liu and Ting Liu .....	1381
<i>Learning Dialog Policies from Weak Demonstrations</i> Gabriel Gordon-Hall, Philip John Gorinski and Shay B. Cohen .....	1394
<i>MuTual: A Dataset for Multi-Turn Dialogue Reasoning</i> Leyang Cui, Yu Wu, Shujie Liu, Yue Zhang and Ming Zhou .....	1406
<i>You Impress Me: Dialogue Generation via Mutual Persona Perception</i> Qian Liu, Yihong Chen, Bei Chen, Jian-Guang LOU, Zixuan Chen, Bin Zhou and Dongmei Zhang	1417
<i>Bridging Anaphora Resolution as Question Answering</i> Yufang Hou .....	1428
<i>Dialogue Coherence Assessment Without Explicit Dialogue Act Labels</i> Mohsen Mesgar, Sebastian Bückner and Iryna Gurevych .....	1439
<i>Fast and Accurate Non-Projective Dependency Tree Linearization</i> Xiang Yu, Simon Tannert, Ngoc Thang Vu and Jonas Kuhn .....	1451
<i>Semantic Graphs for Generating Deep Questions</i> Liangming Pan, Yuxi Xie, Yansong Feng, Tat-Seng Chua and Min-Yen Kan .....	1463
<i>A Novel Cascade Binary Tagging Framework for Relational Triple Extraction</i> Zhepei Wei, Jianlin Su, Yue Wang, Yuan Tian and Yi Chang .....	1476
<i>In Layman’s Terms: Semi-Open Relation Extraction from Scientific Texts</i> Ruben Kruiper, Julian Vincent, Jessica Chen-Burger, Marc Desmulliez and Ioannis Konstas ..	1489
<i>NAT: Noise-Aware Training for Robust Neural Sequence Labeling</i> Marcin Namysl, Sven Behnke and Joachim Köhler .....	1501
<i>Named Entity Recognition without Labelled Data: A Weak Supervision Approach</i> Pierre Lison, Jeremy Barnes, Aliaksandr Hubin and Samia Touileb .....	1518

<i>Probing Linguistic Features of Sentence-Level Representations in Relation Extraction</i> Christoph Alt, Aleksandra Gabryszak and Leonhard Hennig .....	1534
<i>Reasoning with Latent Structure Refinement for Document-Level Relation Extraction</i> Guoshun Nan, Zhijiang Guo, Ivan Sekulic and Wei Lu .....	1546
<i>TACRED Revisited: A Thorough Evaluation of the TACRED Relation Extraction Task</i> Christoph Alt, Aleksandra Gabryszak and Leonhard Hennig .....	1558
<i>Bilingual Dictionary Based Neural Machine Translation without Using Parallel Sentences</i> Xiangyu Duan, Baijun Ji, Hao Jia, Min Tan, Min Zhang, Boxing Chen, Weihua Luo and Yue Zhang 1570	
<i>Boosting Neural Machine Translation with Similar Translations</i> Jitao XU, Josep Crego and Jean Senellart .....	1580
<i>Character-Level Translation with Self-attention</i> Yingqiang Gao, Nikola I. Nikolov, Yuhuang Hu and Richard H.R. Hahnloser .....	1591
<i>End-to-End Neural Word Alignment Outperforms GIZA++</i> Thomas Zenkel, Joern Wuebker and John DeNero .....	1605
<i>Enhancing Machine Translation with Dependency-Aware Self-Attention</i> Emanuele Bugliarello and Naoaki Okazaki .....	1618
<i>Improving Massively Multilingual Neural Machine Translation and Zero-Shot Translation</i> Biao Zhang, Philip Williams, Ivan Titov and Rico Sennrich .....	1628
<i>It's Easier to Translate out of English than into it: Measuring Neural Translation Difficulty by Cross-Mutual Information</i> Emanuele Bugliarello, Sabrina J. Mielke, Antonios Anastasopoulos, Ryan Cotterell and Naoaki Okazaki .....	1640
<i>Language-aware Interlingua for Multilingual Neural Machine Translation</i> Changfeng Zhu, Heng Yu, Shanbo Cheng and Weihua Luo .....	1650
<i>On the Limitations of Cross-lingual Encoders as Exposed by Reference-Free Machine Translation Evaluation</i> Wei Zhao, Goran Glavaš, Maxime Peyrard, Yang Gao, Robert West and Steffen Eger .....	1656
<i>Parallel Sentence Mining by Constrained Decoding</i> Pinzhen Chen, Nikolay Bogoychev, Kenneth Heafield and Faheem Kirefu .....	1672
<i>Self-Attention with Cross-Lingual Position Representation</i> Liang Ding, Longyue Wang and Dacheng Tao .....	1679
<i>“You Sound Just Like Your Father” Commercial Machine Translation Systems Include Stylistic Biases</i> Dirk Hovy, Federico Bianchi and Tommaso Fornaciari .....	1686
<i>MMPE: A Multi-Modal Interface for Post-Editing Machine Translation</i> Nico Herbig, Tim Düwel, Santanu Pal, Kalliopi Meladaki, Mahsa Monshizadeh, Antonio Krüger and Josef van Genabith .....	1691
<i>A Monolingual Approach to Contextualized Word Embeddings for Mid-Resource Languages</i> Pedro Javier Ortiz Suárez, Laurent Romary and Benoît Sagot .....	1703

<i>Will-They-Won't-They: A Very Large Dataset for Stance Detection on Twitter</i> Costanza Conforti, Jakob Berndt, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd and Nigel Collier .....	1715
<i>A Systematic Assessment of Syntactic Generalization in Neural Language Models</i> Jennifer Hu, Jon Gauthier, Peng Qian, Ethan Wilcox and Roger Levy .....	1725
<i>Inflecting When There's No Majority: Limitations of Encoder-Decoder Neural Networks as Cognitive Models for German Plurals</i> Kate McCurdy, Sharon Goldwater and Adam Lopez .....	1745
<i>Overestimation of Syntactic Representation in Neural Language Models</i> Jordan Kodner and Nitish Gupta .....	1757
<i>Suspense in Short Stories is Predicted By Uncertainty Reduction over Neural Story Representation</i> David Wilmot and Frank Keller .....	1763
<i>You Don't Have Time to Read This: An Exploration of Document Reading Time Prediction</i> Orion Weller, Jordan Hildebrandt, Ilya Reznik, Christopher Challis, E. Shannon Tass, Quinn Snell and Kevin Seppi .....	1789
<i>A Generative Model for Joint Natural Language Understanding and Generation</i> Bo-Hsiang Tseng, Jianpeng Cheng, Yimai Fang and David Vandyke .....	1795
<i>Automatic Detection of Generated Text is Easiest when Humans are Fooled</i> Daphne Ippolito, Daniel Duckworth, Chris Callison-Burch and Douglas Eck .....	1808
<i>Multi-Domain Neural Machine Translation with Word-Level Adaptive Layer-wise Domain Mixing</i> Haoming Jiang, Chen Liang, Chong Wang and Tuo Zhao .....	1823
<i>Conversational Graph Grounded Policy Learning for Open-Domain Conversation Generation</i> Jun Xu, Haifeng Wang, Zheng-Yu Niu, Hua Wu, Wanxiang Che and Ting Liu .....	1835
<i>GPT-too: A Language-Model-First Approach for AMR-to-Text Generation</i> Manuel Mager, Ramón Fernandez Astudillo, Tahira Naseem, Md Arafat Sultan, Young-Suk Lee, Radu Florian and Salim Roukos .....	1846
<i>Learning to Update Natural Language Comments Based on Code Changes</i> Sheena Panthaplackel, Pengyu Nie, Milos Gligoric, Junyi Jessy Li and Raymond Mooney ...	1853
<i>Politeness Transfer: A Tag and Generate Approach</i> Aman Madaan, Amrith Setlur, Tanmay Parekh, Barnabas Poczos, Graham Neubig, Yiming Yang, Ruslan Salakhutdinov, Alan W Black and Shrimai Prabhunoye .....	1869
<i>BPE-Dropout: Simple and Effective Subword Regularization</i> Ivan Provilkov, Dmitrii Emelianenko and Elena Voita .....	1882
<i>Improving Non-autoregressive Neural Machine Translation with Monolingual Data</i> Jiawei Zhou and Phillip Keung .....	1893
<i>Attend to Medical Ontologies: Content Selection for Clinical Abstractive Summarization</i> Sajad Sotudeh Gharebagh, Nazli Goharian and Ross Filice .....	1899
<i>On Faithfulness and Factuality in Abstractive Summarization</i> Joshua Maynez, Shashi Narayan, Bernd Bohnet and Ryan McDonald .....	1906

<i>Screenplay Summarization Using Latent Narrative Structure</i>	
Pinelopi Papalampidi, Frank Keller, Lea Frermann and Mirella Lapata . . . . .	1920
<i>Unsupervised Opinion Summarization with Noising and Denoising</i>	
Reinald Kim Amplayo and Mirella Lapata . . . . .	1934
<i>A Tale of Two Perplexities: Sensitivity of Neural Language Models to Lexical Retrieval Deficits in Dementia of the Alzheimer's Type</i>	
Trevor Cohen and Serguei Pakhomov . . . . .	1946
<i>Probing Linguistic Systematicity</i>	
Emily Goodwin, Koustuv Sinha and Timothy J. O'Donnell . . . . .	1958
<i>Recollection versus Imagination: Exploring Human Memory and Cognition via Neural Language Models</i>	
Maarten Sap, Eric horvitz, Yejin Choi, Noah A. Smith and James Pennebaker . . . . .	1970
<i>Recurrent Neural Network Language Models Always Learn English-Like Relative Clause Attachment</i>	
Forrest Davis and Marten van Schijndel . . . . .	1979
<i>Speakers enhance contextually confusable words</i>	
Eric Meinhardt, Eric Bakovic and Leon Bergen . . . . .	1991
<i>What determines the order of adjectives in English? Comparing efficiency-based theories using dependency treebanks</i>	
Richard Futrell, William Dyer and Greg Scontras . . . . .	2003
<i>"None of the Above": Measure Uncertainty in Dialog Response Retrieval</i>	
Yulan Feng, Shikib Mehri, Maxine Eskenazi and Tiancheng Zhao . . . . .	2013
<i>Can You Put it All Together: Evaluating Conversational Agents' Ability to Blend Skills</i>	
Eric Michael Smith, Mary Williamson, Kurt Shuster, Jason Weston and Y-Lan Boureau . . . . .	2021
<i>Grounded Conversation Generation as Guided Traverses in Commonsense Knowledge Graphs</i>	
Houyu Zhang, Zhenghao Liu, Chenyan Xiong and Zhiyuan Liu . . . . .	2031
<i>Negative Training for Neural Dialogue Response Generation</i>	
Tianxing He and James Glass . . . . .	2044
<i>Recursive Template-based Frame Generation for Task Oriented Dialog</i>	
Rashmi Gangadharaiah and Balakrishnan Narayanaswamy . . . . .	2059
<i>Speak to your Parser: Interactive Text-to-SQL with Natural Language Feedback</i>	
Ahmed Elgohary, saghar Hosseini and Ahmed Hassan Awadallah . . . . .	2065
<i>Calibrating Structured Output Predictors for Natural Language Processing</i>	
Abhyuday Jagannatha and hong yu . . . . .	2078
<i>Active Imitation Learning with Noisy Guidance</i>	
Kianté Brantley, Hal Daumé III and Amr Sharaf . . . . .	2093
<i>ExpBERT: Representation Engineering with Natural Language Explanations</i>	
Shikhar Murty, Pang Wei Koh and Percy Liang . . . . .	2106

<i>GAN-BERT: Generative Adversarial Learning for Robust Text Classification with a Bunch of Labeled Examples</i>	
Danilo Croce, Giuseppe Castellucci and Roberto Basili .....	2114
<i>Generalizing Natural Language Analysis through Span-relation Representations</i>	
Zhengbao Jiang, Wei Xu, Jun Araki and Graham Neubig .....	2120
<i>Learning to Contextually Aggregate Multi-Source Supervision for Sequence Labeling</i>	
Ouyu Lan, Xiao Huang, Bill Yuchen Lin, He Jiang, Liyuan Liu and Xiang Ren .....	2134
<i>MixText: Linguistically-Informed Interpolation of Hidden Space for Semi-Supervised Text Classification</i>	
Jiaao Chen, Zichao Yang and Diyi Yang .....	2147
<i>MobileBERT: a Compact Task-Agnostic BERT for Resource-Limited Devices</i>	
Zhiqing Sun, Hongkun Yu, Xiaodan Song, Renjie Liu, Yiming Yang and Denny Zhou .....	2158
<i>On Importance Sampling-Based Evaluation of Latent Language Models</i>	
Robert L Logan IV, Matt Gardner and Sameer Singh .....	2171
<i>SMART: Robust and Efficient Fine-Tuning for Pre-trained Natural Language Models through Principled Regularized Optimization</i>	
Haoming Jiang, Pengcheng He, Weizhu Chen, Xiaodong Liu, Jianfeng Gao and Tuo Zhao ...	2177
<i>Stolen Probability: A Structural Weakness of Neural Language Models</i>	
David Demeter, Gregory Kimmel and Doug Downey .....	2191
<i>Taxonomy Construction of Unseen Domains via Graph-based Cross-Domain Knowledge Transfer</i>	
Chao Shang, Sarthak Dash, Md. Faisal Mahbub Chowdhury, Nandana Mihindukulasooriya and Alfio Gliozzo .....	2198
<i>To Pretrain or Not to Pretrain: Examining the Benefits of Pretraining on Resource Rich Tasks</i>	
Sinong Wang, Madian Khabisa and Hao Ma .....	2209
<i>Why Overfitting Isn't Always Bad: Retrofitting Cross-Lingual Word Embeddings to Dictionaries</i>	
Mozhi Zhang, Yoshinari Fujinuma, Michael J. Paul and Jordan Boyd-Graber .....	2214
<i>XtremeDistil: Multi-stage Distillation for Massive Multilingual Models</i>	
Subhabrata Mukherjee and Ahmed Hassan Awadallah .....	2221
<i>A Girl Has A Name: Detecting Authorship Obfuscation</i>	
Asad Mahmood, Zubair Shafiq and Padmini Srinivasan .....	2235
<i>DeeBERT: Dynamic Early Exiting for Accelerating BERT Inference</i>	
Ji Xin, Raphael Tang, Jaejun Lee, Yaoliang Yu and Jimmy Lin .....	2246
<i>Efficient Strategies for Hierarchical Text Classification: External Knowledge and Auxiliary Tasks</i>	
Kervy Rivas Rojas, Gina Bustamante, Arturo Oncevay and Marco Antonio Sobrevilla Cabezudo	2252
<i>Investigating the effect of auxiliary objectives for the automated grading of learner English speech transcriptions</i>	
Hannah Craighead, Andrew Caines, Paula Buttery and Helen Yannakoudakis .....	2258
<i>SPECTER: Document-level Representation Learning using Citation-informed Transformers</i>	
Arman Cohan, Sergey Feldman, Iz Beltagy, Doug Downey and Daniel Weld .....	2270

<i>Semantic Scaffolds for Pseudocode-to-Code Generation</i>	
Ruiqi Zhong, Mitchell Stern and Dan Klein .....	2283
<i>Can We Predict New Facts with Open Knowledge Graph Embeddings? A Benchmark for Open Link Prediction</i>	
Samuel Broscheit, Kiril Gashteovski, Yanjie Wang and Rainer Gemulla .....	2296
<i>INFOTABS: Inference on Tables as Semi-structured Data</i>	
Vivek Gupta, Maitrey Mehta, Pegah Nokhiz and Vivek Srikumar .....	2309
<i>Interactive Machine Comprehension with Information Seeking Agents</i>	
Xingdi Yuan, Jie Fu, Marc-Alexandre Côté, Yi Tay, Chris Pal and Adam Trischler .....	2325
<i>Syntactic Data Augmentation Increases Robustness to Inference Heuristics</i>	
Junghyun Min, R. Thomas McCoy, Dipanjan Das, Emily Pitler and Tal Linzen .....	2339
<i>Improved Speech Representations with Multi-Target Autoregressive Predictive Coding</i>	
Yu-An Chung and James Glass .....	2353
<i>Integrating Multimodal Information in Large Pretrained Transformers</i>	
Wasifur Rahman, Md Kamrul Hasan, Sangwu Lee, AmirAli Bagher Zadeh, Chengfeng Mao, Louis-Philippe Morency and Ehsan Hoque .....	2359
<i>MultiQT: Multimodal learning for real-time question tracking in speech</i>	
Jakob D. Havtorn, Jan Latko, Joakim Edin, Lars Maaløe, Lasse Borgholt, Lorenzo Belgrano, Nicolai Jacobsen, Regitze Sdun and Željko Agić .....	2370
<i>Multimodal and Multiresolution Speech Recognition with Transformers</i>	
Georgios Paraskevopoulos, Srinivas Parthasarathy, Aparna Khare and Shiva Sundaram .....	2381
<i>Phone Features Improve Speech Translation</i>	
Elizabeth Salesky and Alan W Black .....	2388
<i>Grounding Conversations with Improvised Dialogues</i>	
Hyundong Cho and Jonathan May .....	2398
<i>Image-Chat: Engaging Grounded Conversations</i>	
Kurt Shuster, Samuel Humeau, Antoine Bordes and Jason Weston .....	2414
<i>Learning an Unreferenced Metric for Online Dialogue Evaluation</i>	
Koustuv Sinha, Prasanna Parthasarathi, Jasmine Wang, Ryan Lowe, William L. Hamilton and Joelle Pineau .....	2430
<i>Neural Generation of Dialogue Response Timings</i>	
Matthew Roddy and Naomi Harte .....	2442
<i>The Dialogue Dodecathlon: Open-Domain Knowledge and Image Grounded Conversational Agents</i>	
Kurt Shuster, Da JU, Stephen Roller, Emily Dinan, Y-Lan Boureau and Jason Weston .....	2453
<i>Automatic Poetry Generation from Prosaic Text</i>	
Tim Van de Cruys .....	2471
<i>Bridging the Structural Gap Between Encoding and Decoding for Data-To-Text Generation</i>	
Chao Zhao, Marilyn Walker and Snigdha Chaturvedi .....	2481



<i>Enabling Language Models to Fill in the Blanks</i>	
Chris Donahue, Mina Lee and Percy Liang .....	2492
<i>INSET: Sentence Infilling with INter-SEntential Transformer</i>	
Yichen Huang, Yizhe Zhang, Oussama Elachqar and Yu Cheng .....	2502
<i>Improving Adversarial Text Generation by Modeling the Distant Future</i>	
Ruiyi Zhang, Changyou Chen, Zhe Gan, Wenlin Wang, Dinghan Shen, Guoyin Wang, Zheng Wen and Lawrence Carin .....	2516
<i>Simple and Effective Retrieve-Edit-Rerank Text Generation</i>	
Nabil Hossain, Marjan Ghazvininejad and Luke Zettlemoyer .....	2532
<i>BabyWalk: Going Farther in Vision-and-Language Navigation by Taking Baby Steps</i>	
Wang Zhu, Hexiang Hu, Jiacheng Chen, Zhiwei Deng, Vihan Jain, Eugene Ie and Fei Sha ...	2539
<i>Cross-media Structured Common Space for Multimedia Event Extraction</i>	
Manling Li, Alireza Zareian, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji and Shih-Fu Chang	2557
<i>Learning to Segment Actions from Observation and Narration</i>	
Daniel Fried, Jean-Baptiste Alayrac, Phil Blunsom, Chris Dyer, Stephen Clark and Aida Nematzadeh .....	2569
<i>Learning to execute instructions in a Minecraft dialogue</i>	
Prashant Jayannavar, Anjali Narayan-Chen and Julia Hockenmaier .....	2589
<i>MART: Memory-Augmented Recurrent Transformer for Coherent Video Paragraph Captioning</i>	
Jie Lei, Liwei Wang, Yelong Shen, Dong Yu, Tamara Berg and Mohit Bansal .....	2603
<i>What is Learned in Visually Grounded Neural Syntax Acquisition</i>	
Noriyuki Kojima, Hadar Averbuch-Elor, Alexander Rush and Yoav Artzi .....	2615
<i>A Batch Normalized Inference Network Keeps the KL Vanishing Away</i>	
Qile Zhu, Wei Bi, Xiaojiang Liu, Xiyao Ma, Xiaolin Li and Dapeng Wu .....	2636
<i>Contextual Embeddings: When Are They Worth It?</i>	
Simran Arora, Avner May, Jian Zhang and Christopher Ré .....	2650
<i>Interactive Classification by Asking Informative Questions</i>	
Lili Yu, Howard Chen, Sida I. Wang, Tao Lei and Yoav Artzi .....	2664
<i>Knowledge Graph Embedding Compression</i>	
Mrinmaya Sachan .....	2681
<i>Low Resource Sequence Tagging using Sentence Reconstruction</i>	
Tal Perl, Sriram Chaudhury and Raja Giryes .....	2692
<i>Masked Language Model Scoring</i>	
Julian Salazar, Davis Liang, Toan Q. Nguyen and Katrin Kirchhoff .....	2699
<i>Orthogonal Relation Transforms with Graph Context Modeling for Knowledge Graph Embedding</i>	
Yun Tang, Jing Huang, Guangtao Wang, Xiaodong He and Bowen Zhou .....	2713
<i>Posterior Calibrated Training on Sentence Classification Tasks</i>	
Taehee Jung, Dongyeop Kang, Hua Cheng, Lucas Mentch and Thomas Schaaf .....	2723



<i>Posterior Control of Blackbox Generation</i>	
Xiang Lisa Li and Alexander Rush .....	2731
<i>Pretrained Transformers Improve Out-of-Distribution Robustness</i>	
Dan Hendrycks, Xiaoyuan Liu, Eric Wallace, Adam Dziedziec, Rishabh Krishnan and Dawn Song	
2744	
<i>Robust Encodings: A Framework for Combating Adversarial Typos</i>	
Erik Jones, Robin Jia, Aditi Raghunathan and Percy Liang .....	2752
<i>Showing Your Work Doesn't Always Work</i>	
Raphael Tang, Jaejun Lee, Ji Xin, Xinyu Liu, Yaoliang Yu and Jimmy Lin .....	2766
<i>Span Selection Pre-training for Question Answering</i>	
Michael Glass, Alfio Gliozzo, Rishabh Chakravarti, Anthony Ferritto, Lin Pan, G P Shrivatsa Bhargav, Dinesh Garg and Avi Sil .....	2773
<i>Topological Sort for Sentence Ordering</i>	
Shrimai Prabhumoye, Ruslan Salakhutdinov and Alan W Black .....	2783
<i>Weight Poisoning Attacks on Pretrained Models</i>	
Keita Kurita, Paul Michel and Graham Neubig .....	2793
<i>schuBERT: Optimizing Elements of BERT</i>	
Ashish Khetan and Zohar Karnin .....	2807
<i>ENGINE: Energy-Based Inference Networks for Non-Autoregressive Machine Translation</i>	
Lifu Tu, Richard Yuanzhe Pang, Sam Wiseman and Kevin Gimpel .....	2819
<i>Leveraging Monolingual Data with Self-Supervision for Multilingual Neural Machine Translation</i>	
Aditya Siddhant, Ankur Bapna, Yuan Cao, Orhan Firat, Mia Chen, Sneha Kudugunta, Naveen Arivazhagan and Yonghui Wu .....	2827
<i>On The Evaluation of Machine Translation Systems Trained With Back-Translation</i>	
Sergey Edunov, Myle Ott, Marc'Aurelio Ranzato and Michael Auli .....	2836
<i>Simultaneous Translation Policies: From Fixed to Adaptive</i>	
Baigong Zheng, Kaibo Liu, Renjie Zheng, Mingbo Ma, Hairong Liu and Liang Huang .....	2847
<i>Breaking Through the 80% Glass Ceiling: Raising the State of the Art in Word Sense Disambiguation by Incorporating Knowledge Graph Information</i>	
Michele Bevilacqua and Roberto Navigli .....	2854
<i>Glyph2Vec: Learning Chinese Out-of-Vocabulary Word Embedding from Glyphs</i>	
Hong-You Chen, SZ-HAN YU and Shou-de Lin .....	2865
<i>Multidirectional Associative Optimization of Function-Specific Word Representations</i>	
Daniela Gerz, Ivan Vulić, Marek Rei, Roi Reichart and Anna Korhonen .....	2872
<i>Predicting Degrees of Technicality in Automatic Terminology Extraction</i>	
Anna HäTTY, Dominik Schlechtweg, Michael Dorna and Sabine Schulte im Walde .....	2883
<i>Verbal Multiword Expressions for Identification of Metaphor</i>	
Omid Rohanian, Marek Rei, Shiva Taslimipoor and Le An Ha .....	2890

<i>Gender Bias in Multilingual Embeddings and Cross-Lingual Transfer</i>	
Jieyu Zhao, Subhabrata Mukherjee, saghar Hosseini, Kai-Wei Chang and Ahmed Hassan Awadallah	
	2896
<i>Give Me Convenience and Give Her Death: Who Should Decide What Uses of NLP are Appropriate, and on What Basis?</i>	
kobi leins, Jey Han Lau and Timothy Baldwin	2908
<i>Is Your Classifier Actually Biased? Measuring Fairness under Uncertainty with Bernstein Bounds</i>	
Kawin Ethayarajh	2914
<i>It's Morphin' Time! Combating Linguistic Discrimination with Inflectional Perturbations</i>	
Samson Tan, Shafiq Joty, Min-Yen Kan and Richard Socher	2920
<i>Mitigating Gender Bias Amplification in Distribution by Posterior Regularization</i>	
Shengyu Jia, Tao Meng, Jieyu Zhao and Kai-Wei Chang	2936
<i>Towards Understanding Gender Bias in Relation Extraction</i>	
Andrew Gaut, Tony Sun, Shirlyn Tang, Yuxin Huang, Jing Qian, Mai ElSherief, Jieyu Zhao, Diba Mirza, Elizabeth Belding, Kai-Wei Chang and William Yang Wang	2943
<i>A Probabilistic Generative Model for Typographical Analysis of Early Modern Printing</i>	
Kartik Goyal, Chris Dyer, Christopher Warren, Maxwell G'Sell and Taylor Berg-Kirkpatrick	2954
<i>Attentive Pooling with Learnable Norms for Text Representation</i>	
Chuhan Wu, Fangzhao Wu, Tao Qi, Xiaohui Cui and Yongfeng Huang	2961
<i>Estimating the influence of auxiliary tasks for multi-task learning of sequence tagging tasks</i>	
Fynn Schröder and Chris Biemann	2971
<i>How Does Selective Mechanism Improve Self-Attention Networks?</i>	
Xinwei Geng, Longyue Wang, Xing Wang, Bing Qin, Ting Liu and Zhaopeng Tu	2986
<i>Improving Transformer Models by Reordering their Sublayers</i>	
Ofir Press, Noah A. Smith and Omer Levy	2996
<i>Single Model Ensemble using Pseudo-Tags and Distinct Vectors</i>	
Ryosuke Kuwabara, Jun Suzuki and Hideki Nakayama	3006
<i>Zero-shot Text Classification via Reinforced Self-training</i>	
Zhiquan Ye, Yuxia Geng, Jiaoyan Chen, Jingmin Chen, Xiaoxiao Xu, SuHang Zheng, Feng Wang, Jun Zhang and Huajun Chen	3014
<i>A Novel Graph-based Multi-modal Fusion Encoder for Neural Machine Translation</i>	
Yongjing Yin, Fandong Meng, Jinsong Su, Chulun Zhou, Zhengyuan Yang, Jie Zhou and Jiebo Luo	3025
<i>A Relaxed Matching Procedure for Unsupervised BLI</i>	
Xu Zhao, Zihao Wang, Yong Zhang and Hao Wu	3036
<i>Dynamic Programming Encoding for Subword Segmentation in Neural Machine Translation</i>	
Xuanli He, Gholamreza Haffari and Mohammad Norouzi	3042
<i>Geometry-aware domain adaptation for unsupervised alignment of word embeddings</i>	
Pratik Jawanpuria, Mayank Meghwanshi and Bamdev Mishra	3052

<i>Learning to Recover from Multi-Modality Errors for Non-Autoregressive Neural Machine Translation</i> Qiu Ran, Yankai Lin, Peng Li and Jie Zhou .....	3059
<i>On the Inference Calibration of Neural Machine Translation</i> Shuo Wang, Zhaopeng Tu, Shuming Shi and Yang Liu .....	3070
<i>Camouflaged Chinese Spam Content Detection with Semi-supervised Generative Active Learning</i> Zhuoren Jiang, Zhe Gao, Yu Duan, Yangyang Kang, Changlong Sun, Qiong Zhang and Xiaozhong Liu .....	3080
<i>Distinguish Confusing Law Articles for Legal Judgment Prediction</i> Nuo Xu, Pinghui Wang, Long Chen, Li Pan, Xiaoyan Wang and Junzhou Zhao .....	3086
<i>Hiring Now: A Skill-Aware Multi-Attention Model for Job Posting Generation</i> Liting Liu, Jie Liu, Wenzheng Zhang, Ziming Chi, Wenxuan Shi and Yalou Huang .....	3096
<i>HyperCore: Hyperbolic and Co-graph Representation for Automatic ICD Coding</i> Pengfei Cao, Yubo Chen, Kang Liu, Jun Zhao, Shengping Liu and Weifeng Chong .....	3105
<i>Hyperbolic Capsule Networks for Multi-Label Classification</i> Boli Chen, Xin Huang, Lin Xiao and Liping Jing .....	3115
<i>Improving Segmentation for Technical Support Problems</i> Kushal Chauhan and Abhirut Gupta .....	3125
<i>MOOCCube: A Large-scale Data Repository for NLP Applications in MOOCs</i> Jifan Yu, Gan Luo, Tong Xiao, Qingyang Zhong, Yuquan Wang, wenzheng feng, Junyi Luo, Chenyu Wang, Lei Hou, Juanzi Li, Zhiyuan Liu and Jie Tang .....	3135
<i>Towards Interpretable Clinical Diagnosis with Bayesian Network Ensembles Stacked on Entity-Aware CNNs</i> Jun Chen, Xiaoya Dai, Quan Yuan, Chao Lu and Haifeng Huang .....	3143
<i>Analyzing the Persuasive Effect of Style in News Editorial Argumentation</i> Roxanne El Baff, Henning Wachsmuth, Khalid Al Khatib and Benno Stein .....	3154
<i>ECPE-2D: Emotion-Cause Pair Extraction based on Joint Two-Dimensional Representation, Interaction and Prediction</i> Zixiang Ding, Rui Xia and Jianfei Yu .....	3161
<i>Effective Inter-Clause Modeling for End-to-End Emotion-Cause Pair Extraction</i> Penghui Wei, Jiahao Zhao and Wenji Mao .....	3171
<i>Embarrassingly Simple Unsupervised Aspect Extraction</i> Stéphan Tulkens and Andreas van Cranenburgh .....	3182
<i>Enhancing Cross-target Stance Detection with Transferable Semantic-Emotion Knowledge</i> Bowen Zhang, Min Yang, Xutao Li, Yunming Ye, Xiaofei Xu and Kuai Dai .....	3188
<i>KinGDOM: Knowledge-Guided DOMain Adaptation for Sentiment Analysis</i> Deepanway Ghosal, Devamanyu Hazarika, Abhinaba Roy, Navonil Majumder, Rada Mihalcea and Soujanya Poria .....	3198
<i>Modelling Context and Syntactical Features for Aspect-based Sentiment Analysis</i> Minh Hieu Phan and Philip O. Ogunbona .....	3211

<i>Parallel Data Augmentation for Formality Style Transfer</i>	
Yi Zhang, Tao Ge and Xu SUN .....	3221
<i>Relational Graph Attention Network for Aspect-based Sentiment Analysis</i>	
Kai Wang, Weizhou Shen, Yunyi Yang, Xiaojun Quan and Rui Wang .....	3229
<i>SpanMtl: A Span-based Multi-Task Learning Framework for Pair-wise Aspect and Opinion Terms Extraction</i>	
He Zhao, Longtao Huang, Rong Zhang, Quan Lu and hui xue .....	3239
<i>Syntax-Aware Opinion Role Labeling with Dependency Graph Convolutional Networks</i>	
Bo Zhang, Yue Zhang, Rui Wang, Zhenghua Li and Min Zhang .....	3249
<i>Towards Better Non-Tree Argument Mining: Proposition-Level Biaffine Parsing with Task-Specific Parameterization</i>	
Gaku Morio, Hiroaki Ozaki, Terufumi Morishita, Yuta Koreeda and Kohsuke Yanai .....	3259
<i>A Span-based Linearization for Constituent Trees</i>	
Yang Wei, Yuanbin Wu and Man Lan .....	3267
<i>An Empirical Comparison of Unsupervised Constituency Parsing Methods</i>	
Jun Li, Yifan Cao, Jiong Cai, Yong Jiang and Kewei Tu .....	3278
<i>Efficient Constituency Parsing by Pointing</i>	
Thanh-Tung Nguyen, Xuan-Phi Nguyen, Shafiq Joty and Xiaoli Li .....	3284
<i>Efficient Second-Order TreeCRF for Neural Dependency Parsing</i>	
Yu Zhang, Zhenghua Li and Min Zhang .....	3295
<i>Representations of Syntax [MASK] Useful: Effects of Constituency and Dependency Structure in Recursive LSTMs</i>	
Michael Lepori, Tal Linzen and R. Thomas McCoy .....	3306
<i>Structure-Level Knowledge Distillation For Multilingual Sequence Labeling</i>	
Xinyu Wang, Yong Jiang, Nguyen Bach, Tao Wang, Fei Huang and Kewei Tu .....	3317
<i>Dynamic Online Conversation Recommendation</i>	
Xingshan Zeng, Jing Li, Lu Wang, Zhiming Mao and Kam-Fai Wong .....	3331
<i>Improving Multimodal Named Entity Recognition via Entity Span Detection with Unified Multimodal Transformer</i>	
Jianfei Yu, Jing Jiang, Li Yang and Rui Xia .....	3342
<i>Stock Embeddings Acquired from News Articles and Price History, and an Application to Portfolio Optimization</i>	
Xin Du and Kumiko Tanaka-Ishii .....	3353
<i>What Was Written vs. Who Read It: News Media Profiling Using Text Analysis and Social Media Context</i>	
Ramy Baly, Georgi Karadzhov, Jisun An, Haewoon Kwak, Yoan Dinkov, Ahmed Ali, James Glass and Preslav Nakov .....	3364
<i>An Analysis of the Utility of Explicit Negative Examples to Improve the Syntactic Abilities of Neural Language Models</i>	
Hiroshi Noji and Hiroya Takamura .....	3375

<i>On the Robustness of Language Encoders against Grammatical Errors</i> Fan Yin, Quanyu Long, Tao Meng and Kai-Wei Chang .....	3386
<i>Roles and Utilization of Attention Heads in Transformer-based Neural Language Models</i> Jae-young Jo and Sung-Hyon Myaeng .....	3404
<i>Understanding Attention for Text Classification</i> Xiaobing Sun and Wei Lu .....	3418
<i>A Relational Memory-based Embedding Model for Triple Classification and Search Personalization</i> Dai Quoc Nguyen, Tu Nguyen and Dinh Phung .....	3429
<i>Do you have the right scissors? Tailoring Pre-trained Language Models via Monte-Carlo Methods</i> Ning Miao, Yuxuan Song, Hao Zhou and Lei Li .....	3436
<i>Enhancing Pre-trained Chinese Character Representation with Word-aligned Attention</i> Yanzeng Li, Bowen Yu, Xue Mengge and Tingwen Liu .....	3442
<i>On the Encoder-Decoder Incompatibility in Variational Text Modeling and Beyond</i> Chen Wu, Prince Zizhuang Wang and William Yang Wang .....	3449
<i>SAFER: A Structure-free Approach for Certified Robustness to Adversarial Word Substitutions</i> Mao Ye, Chengyue Gong and Qiang Liu .....	3465
<i>A Graph-based Coarse-to-fine Method for Unsupervised Bilingual Lexicon Induction</i> Shuo Ren, Shujie Liu, Ming Zhou and Shuai Ma .....	3476
<i>A Reinforced Generation of Adversarial Examples for Neural Machine Translation</i> wei zou, Shujian Huang, Jun Xie, Xinyu Dai and Jiajun CHEN .....	3486
<i>A Retrieve-and-Rewrite Initialization Method for Unsupervised Machine Translation</i> Shuo Ren, Yu Wu, Shujie Liu, Ming Zhou and Shuai Ma .....	3498
<i>A Simple and Effective Unified Encoder for Document-Level Machine Translation</i> Shuming Ma, Dongdong Zhang and Ming Zhou .....	3505
<i>Does Multi-Encoder Help? A Case Study on Context-Aware Neural Machine Translation</i> Bei Li, Hui Liu, Ziyang Wang, Yufan Jiang, Tong Xiao, Jingbo Zhu, Tongran Liu and changliang li 3512	
<i>Dynamically Adjusting Transformer Batch Size by Monitoring Gradient Direction Change</i> Hongfei Xu, Josef van Genabith, Deyi Xiong and Qiuhui Liu .....	3519
<i>Knowledge Distillation for Multilingual Unsupervised Neural Machine Translation</i> Haipeng Sun, Rui Wang, Kehai Chen, Masao Utiyama, Eiichiro Sumita and Tiejun Zhao ....	3525
<i>Lexically Constrained Neural Machine Translation with Levenshtein Transformer</i> Raymond Hendy Susanto, Shamil Chollampatt and Liling Tan .....	3536
<i>On Exposure Bias, Hallucination and Domain Shift in Neural Machine Translation</i> Chaojun Wang and Rico Sennrich .....	3544
<i>Automatic Machine Translation Evaluation using Source Language Inputs and Cross-lingual Language Model</i> Kosuke Takahashi, Katsuhito Sudoh and Satoshi Nakamura .....	3553

<i>ChartDialogs: Plotting from Natural Language Instructions</i> Yutong Shao and Ndapa Nakashole .....	3559
<i>GLUECoS: An Evaluation Benchmark for Code-Switched NLP</i> Simran Khanuja, Sandipan Dandapat, Anirudh Srinivasan, Sunayana Sitaram and Monojit Choudhury .....	3575
<i>MATINF: A Jointly Labeled Large-Scale Dataset for Classification, Question Answering and Summarization</i> Canwen Xu, Jiaxin Pei, Hongtao Wu, Yiyu Liu and Chenliang Li .....	3586
<i>MIND: A Large-scale Dataset for News Recommendation</i> Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, Chuhan Wu, Tao Qi, Jianxun Lian, Danyang Liu, Xing Xie, Jianfeng Gao, Winnie Wu and Ming Zhou .....	3597
<i>That is a Known Lie: Detecting Previously Fact-Checked Claims</i> Shaden Shaar, Nikolay Babulkov, Giovanni Da San Martino and Preslav Nakov .....	3607
<i>Towards Holistic and Automatic Evaluation of Open-Domain Dialogue Generation</i> Bo Pang, Erik Nijkamp, Wenjuan Han, Linqi Zhou, Yixian Liu and Kewei Tu .....	3619
<i>BiRRE: Learning Bidirectional Residual Relation Embeddings for Supervised Hypernymy Detection</i> Chengyu Wang and XIAOFENG HE .....	3630
<i>Biomedical Entity Representations with Synonym Marginalization</i> Mujeen Sung, Hwisang Jeon, Jinhyuk Lee and Jaewoo Kang .....	3641
<i>Hypernymy Detection for Low-Resource Languages via Meta Learning</i> Changlong Yu, Jialong Han, Haisong Zhang and Wilfred Ng .....	3651
<i>Investigating Word-Class Distributions in Word Vector Spaces</i> Ryohei Sasano and Anna Korhonen .....	3657
<i>Aspect Sentiment Classification with Document-level Sentiment Preference Modeling</i> Xiao Chen, Changlong Sun, Jingjing Wang, Shoushan Li, Luo Si, Min Zhang and Guodong Zhou	3667
<i>Don't Eclipse Your Arts Due to Small Discrepancies: Boundary Repositioning with a Pointer Network for Aspect Extraction</i> Zhenkai Wei, Yu Hong, Bowei Zou, Meng Cheng and Jianmin YAO .....	3678
<i>Relation-Aware Collaborative Learning for Unified Aspect-Based Sentiment Analysis</i> Zhuang Chen and Tiejun Qian .....	3685
<i>SentiBERT: A Transferable Transformer-Based Architecture for Compositional Sentiment Semantics</i> Da Yin, Tao Meng and Kai-Wei Chang .....	3695
<i>Transition-based Directed Graph Construction for Emotion-Cause Pair Extraction</i> Chuang Fan, Chaofa Yuan, Jiachen Du, Lin Gui, Min Yang and Ruifeng Xu .....	3707
<i>CH-SIMS: A Chinese Multimodal Sentiment Analysis Dataset with Fine-grained Annotation of Modality</i> Wenmeng Yu, Hua Xu, Fanyang Meng, Yilin Zhu, Yixiao Ma, Jiele Wu, Jiyun Zou and Kaicheng Yang .....	3718
<i>Curriculum Pre-training for End-to-End Speech Translation</i> Chengyi Wang, Yu Wu, Shujie Liu, Ming Zhou and Zhenglu Yang .....	3728

<i>How Accents Confound: Probing for Accent Information in End-to-End Speech Recognition Systems</i> Archiki Prasad and Preethi Jyothi .....	3739
<i>Improving Disfluency Detection by Self-Training a Self-Attentive Model</i> Paria Jamshid Lou and Mark Johnson .....	3754
<i>Learning Spoken Language Representations with Neural Lattice Language Modeling</i> Chao-Wei Huang and Yun-Nung Chen .....	3764
<i>Meta-Transfer Learning for Code-Switched Speech Recognition</i> Genta Indra Winata, Samuel Cahyawijaya, Zhaojiang Lin, Zihan Liu, Peng Xu and Pascale Fung 3770	
<i>Reasoning with Multimodal Sarcastic Tweets via Modeling Cross-Modality Contrast and Semantic Association</i> Nan Xu, Zhixiong Zeng and Wenji Mao .....	3777
<i>SimulSpeech: End-to-End Simultaneous Speech to Text Translation</i> Yi Ren, Jinglin Liu, Xu Tan, Chen Zhang, Tao QIN, Zhou Zhao and Tie-Yan Liu .....	3787
<i>Towards end-2-end learning for predicting behavior codes from spoken utterances in psychotherapy conversations</i> Karan Singla, Zhuohao Chen, David Atkins and Shrikanth Narayanan .....	3797
<i>Neural Temporal Opinion Modelling for Opinion Prediction on Twitter</i> Lixing Zhu, Yulan He and Deyu Zhou .....	3804
<i>It Takes Two to Lie: One to Lie, and One to Listen</i> Denis Peskov, Benny Cheng, Ahmed Elgohary, Joe Barrow, Cristian Danescu-Niculescu-Mizil and Jordan Boyd-Graber .....	3811
<i>Learning Implicit Text Generation via Feature Matching</i> Inkit Padhi, Pierre Dognin, Ke Bai, Cícero Nogueira dos Santos, Vijil Chenthamarakshan, Youssef Mroueh and Payel Das .....	3855
<i>Two Birds, One Stone: A Simple, Unified Model for Text Generation from Structured and Unstructured Data</i> Hamidreza Shahidi, Ming Li and Jimmy Lin .....	3864
<i>Bayesian Hierarchical Words Representation Learning</i> Oren Barkan, Idan Rejwan, Avi Caciularu and Noam Koenigstein .....	3871
<i>Pre-training Is (Almost) All You Need: An Application to Commonsense Reasoning</i> Alexandre Tamborrino, Nicola Pellicanò, Baptiste Pannier, Pascal Voitot and Louise Naudin .	3878
<i>SEEK: Segmented Embedding of Knowledge Graphs</i> Wentao Xu, Shun Zheng, Liang He, Bin Shao, Jian Yin and Tie-Yan Liu .....	3888
<i>Selecting Backtranslated Data from Multiple Sources for Improved Neural Machine Translation</i> Xabier Soto, Dimitar Shterionov, Alberto Poncelas and Andy Way .....	3898
<i>Successfully Applying the Stabilized Lottery Ticket Hypothesis to the Transformer Architecture</i> Christopher Brix, Parnia Bahar and Hermann Ney .....	3909
<i>A Self-Training Method for Machine Reading Comprehension with Soft Evidence Extraction</i> Yilin Niu, Fangkai Jiao, Mantong Zhou, Ting Yao, jingfang xu and Minlie Huang .....	3916



<i>Graph-to-Tree Learning for Solving Math Word Problems</i>	
Jipeng Zhang, Lei Wang, Roy Ka-Wei Lee, Yi Bin, Yan Wang, Jie Shao and Ee-Peng Lim . . .	3928
<i>An Effectiveness Metric for Ordinal Classification: Formal Properties and Experimental Results</i>	
Enrique Amigo, Julio Gonzalo, Stefano Mizzaro and Jorge Carrillo-de-Albornoz . . . . .	3938
<i>Adaptive Compression of Word Embeddings</i>	
Yeachan Kim, Kang-Min Kim and SangKeun Lee . . . . .	3950
<i>Analysing Lexical Semantic Change with Contextualised Word Representations</i>	
Mario Giulianelli, Marco Del Tredici and Raquel Fernández . . . . .	3960
<i>Autoencoding Keyword Correlation Graph for Document Clustering</i>	
Billy Chiu, Sunil Kumar Sahu, Derek Thomas, Neha Sengupta and Mohammady Mahdy . . . .	3974
<i>Autoencoding Pixies: Amortised Variational Inference with Graph Convolutions for Functional Distributional Semantics</i>	
Guy Emerson . . . . .	3982
<i>BERTRAM: Improved Word Embeddings Have Big Impact on Contextualized Model Performance</i>	
Timo Schick and Hinrich Schütze . . . . .	3996
<i>CluBERT: A Cluster-Based Approach for Learning Sense Distributions in Multiple Languages</i>	
Tommaso Pasini, Federico Scozzafava and Bianca Scarlini . . . . .	4008
<i>Adversarial and Domain-Aware BERT for Cross-Domain Sentiment Analysis</i>	
Chunning Du, Haifeng Sun, Jingyu Wang, Qi Qi and Jianxin Liao . . . . .	4019
<i>From Arguments to Key Points: Towards Automatic Argument Summarization</i>	
Roy Bar-Haim, Lilach Eden, Roni Friedman, Yoav Kantor, Dan Lahav and Noam Slonim . . . .	4029
<i>GoEmotions: A Dataset of Fine-Grained Emotions</i>	
Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade and Sujith Ravi . . . . .	4040
<i>He said "who's gonna take care of your children when you are at ACL?": Reported Sexist Acts are Not Sexist</i>	
Patricia Chiril, Véronique MORICEAU, Farah Benamara, Alda Mari, Gloria Origgi and Marlène Coulomb-Gully . . . . .	4055
<i>SKEP: Sentiment Knowledge Enhanced Pre-training for Sentiment Analysis</i>	
Hao Tian, Can Gao, Xinyan Xiao, Hao Liu, Bolei He, Hua Wu, Haifeng Wang and feng wu . .	4067
<i>Do Neural Language Models Show Preferences for Syntactic Formalisms?</i>	
Artur Kulmizev, Vinit Ravishankar, Mostafa Abdou and Joakim Nivre . . . . .	4077
<i>Enriched In-Order Linearization for Faster Sequence-to-Sequence Constituent Parsing</i>	
Daniel Fernández-González and Carlos Gómez-Rodríguez . . . . .	4092
<i>Exact yet Efficient Graph Parsing, Bi-directional Locality and the Constructivist Hypothesis</i>	
Yajie Ye and Weiwei Sun . . . . .	4100
<i>Max-Margin Incremental CCG Parsing</i>	
Miloš Stanojević and Mark Steedman . . . . .	4111



<i>Neural Reranking for Dependency Parsing: An Evaluation</i>	
Bich-Ngoc Do and Ines Rehbein .....	4123
<i>Demographics Should Not Be the Reason of Toxicity: Mitigating Discrimination in Text Classifications with Instance Weighting</i>	
Guanhua Zhang, Bing Bai, Junqi Zhang, Kun Bai, Conghui Zhu and Tiejun Zhao .....	4134
<i>Analyzing analytical methods: The case of phonology in neural models of spoken language</i>	
Grzegorz Chrupała, Bertrand Higy and Afra Alishahi .....	4146
<i>Make Up Your Mind! Adversarial Generation of Inconsistent Natural Language Explanations</i>	
Oana-Maria Camburu, Brendan Shillingford, Pasquale Minervini, Thomas Lukasiewicz and Phil Blunsom .....	4157
<i>Perturbed Masking: Parameter-free Probing for Analyzing and Interpreting BERT</i>	
Zhiyong Wu, Yun Chen, Ben Kao and Qun Liu .....	4166
<i>Probing for Referential Information in Language Models</i>	
Ionut-Teodor Sorodoc, Kristina Gulordava and Gemma Boleda .....	4177
<i>Quantifying Attention Flow in Transformers</i>	
Samira Abnar and Willem Zuidema .....	4190
<i>Towards Faithfully Interpretable NLP Systems: How Should We Define and Evaluate Faithfulness?</i>	
Alon Jacovi and Yoav Goldberg .....	4198
<i>Towards Transparent and Explainable Attention Models</i>	
Akash Kumar Mohankumar, Preksha Nema, Sharan Narasimhan, Mitesh M. Khapra, Balaji Vasani Srinivasan and Balaraman Ravindran .....	4206
<i>Tchebycheff Procedure for Multi-task Text Classification</i>	
Yuren Mao, Shuang Yun, Weiwei Liu and Bo Du .....	4217
<i>Modeling Word Formation in English–German Neural Machine Translation</i>	
Marion Weller-Di Marco and Alexander Fraser .....	4227
<i>Empowering Active Learning to Jointly Optimize System and User Demands</i>	
Ji-Ung Lee, Christian M. Meyer and Iryna Gurevych .....	4233
<i>Encoder-Decoder Models Can Benefit from Pre-trained Masked Language Models in Grammatical Error Correction</i>	
Masahiro Kaneko, Masato Mita, Shun Kiyono, Jun Suzuki and Kentaro Inui .....	4248
<i>Graph Neural News Recommendation with Unsupervised Preference Disentanglement</i>	
Linmei Hu, Siyong Xu, Chen Li, Cheng Yang, Chuan Shi, Nan Duan, Xing Xie and Ming Zhou .....	4255
<i>Identifying Principals and Accessories in a Complex Case based on the Comprehension of Fact Description</i>	
Yakun Hu, Zhunchen Luo and Wenhan Chao .....	4265
<i>Joint Modelling of Emotion and Abusive Language Detection</i>	
Santhosh Rajamanickam, Pushkar Mishra, Helen Yannakoudakis and Ekaterina Shutova .....	4270

<i>Programming in Natural Language with fuSE: Synthesizing Methods from Spoken Utterances Using Deep Natural Language Understanding</i>	
Sebastian Weigelt, Vanessa Steurer, Tobias Hey and Walter F. Tichy .....	4280
<i>Toxicity Detection: Does Context Really Matter?</i>	
John Pavlopoulos, Jeffrey Sorensen, Lucas Dixon, Nithum Thain and Ion Androutsopoulos ..	4296
<i>AMR Parsing with Latent Structural Information</i>	
Qiji Zhou, Yue Zhang, Donghong Ji and Hao Tang .....	4306
<i>TaPas: Weakly Supervised Table Parsing via Pre-training</i>	
Jonathan Herzig, Pawel Krzysztof Nowak, Thomas Müller, Francesco Piccinno and Julian Eisen- schlos .....	4320
<i>Target Inference in Argument Conclusion Generation</i>	
Milad Alshomary, Shahbaz Syed, Martin Potthast and Henning Wachsmuth .....	4334
<i>Multimodal Transformer for Multimodal Machine Translation</i>	
Shaowei Yao and Xiaojun Wan .....	4346
<i>Sentiment and Emotion help Sarcasm? A Multi-task Learning Framework for Multi-Modal Sarcasm, Sentiment and Emotion Analysis</i>	
Dushyant Singh Chauhan, Dhanush S R, Asif Ekbal and Pushpak Bhattacharyya .....	4351
<i>Towards Emotion-aided Multi-modal Dialogue Act Classification</i>	
Tulika Saha, Aditya Patra, Sriparna Saha and Pushpak Bhattacharyya .....	4361
<i>Analyzing Political Parody in Social Media</i>	
Antonios Maronikolakis, Danae Sánchez Villegas, Daniel Preotiuc-Pietro and Nikolaos Aletras	4373
<i>Masking Actor Information Leads to Fairer Political Claims Detection</i>	
Erenay Dayanik and Sebastian Padó .....	4385
<i>When do Word Embeddings Accurately Reflect Surveys on our Beliefs About People?</i>	
Kenneth Joseph and Jonathan Morgan .....	4392
<i>“Who said it, and Why?” Provenance for Natural Language Claims</i>	
Yi Zhang, Zachary Ives and Dan Roth .....	4416
<i>Compositionality and Generalization In Emergent Languages</i>	
Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux and Marco Baroni	4427
<i>ERASER: A Benchmark to Evaluate Rationalized NLP Models</i>	
Jay DeYoung, Sarthak Jain, Nazneen Fatema Rajani, Eric Lehman, Caiming Xiong, Richard Socher and Byron C. Wallace .....	4443
<i>Learning to Faithfully Rationalize by Construction</i>	
Sarthak Jain, Sarah Wiegrefe, Yuval Pinter and Byron C. Wallace .....	4459
<i>Clinical Reading Comprehension: A Thorough Analysis of the emrQA Dataset</i>	
Xiang Yue, Bernal Jimenez Gutierrez and Huan Sun .....	4474
<i>DeFormer: Decomposing Pre-trained Transformers for Faster Question Answering</i>	
Qingqing Cao, Harsh Trivedi, Aruna Balasubramanian and Niranjan Balasubramanian .....	4487

<i>Improving Multi-hop Question Answering over Knowledge Graphs using Knowledge Base Embeddings</i> Apoorv Saxena, Aditay Tripathi and Partha Talukdar .....	4498
<i>Template-Based Question Generation from Retrieved Sentences for Improved Unsupervised Question Answering</i> Alexander Fabbri, Patrick Ng, Zhiguo Wang, Ramesh Nallapati and Bing Xiang .....	4508
<i>Unsupervised Alignment-based Iterative Evidence Retrieval for Multi-hop Question Answering</i> Vikas Yadav, Steven Bethard and Mihai Surdeanu .....	4514
<i>A Corpus for Large-Scale Phonetic Typology</i> Elizabeth Salesky, Eleanor Chodroff, Tiago Pimentel, Matthew Wiesner, Ryan Cotterell, Alan W Black and Jason Eisner .....	4526
<i>Dscorer: A Fast Evaluation Metric for Discourse Representation Structure Parsing</i> Jiangming Liu, Shay B. Cohen and Mirella Lapata .....	4547
<i>ParaCrawl: Web-Scale Acquisition of Parallel Corpora</i> Marta Bañón, Pinzhen Chen, Barry Haddow, Kenneth Heafield, Hieu Hoang, Miquel Esplà-Gomis, Mikel L. Forcada, Amir Kamran, Faheem Kirefu, Philipp Koehn, Sergio Ortiz Rojas, Leopoldo Pla Sempere, Gema Ramírez-Sánchez, Elsa Sarrías, Marek Strelec, Brian Thompson, William Waites, Dion Wiggins and Jaume Zaragoza .....	4555
<i>Toward Gender-Inclusive Coreference Resolution</i> Yang Trista Cao and Hal Daumé III .....	4568
<i>Human Attention Maps for Text Classification: Do Humans and Neural Networks Focus on the Same Words?</i> Cansu Sen, Thomas Hartvigsen, Biao Yin, Xiangnan Kong and Elke Rundensteiner .....	4596
<i>Information-Theoretic Probing for Linguistic Structure</i> Tiago Pimentel, Josef Valvoda, Rowan Hall Maudslay, Ran Zmigrod, Adina Williams and Ryan Cotterell .....	4609
<i>On the Cross-lingual Transferability of Monolingual Representations</i> Mikel Artetxe, Sebastian Ruder and Dani Yogatama .....	4623
<i>Similarity Analysis of Contextual Word Representation Models</i> John Wu, Yonatan Belinkov, Hassan Sajjad, Nadir Durrani, Fahim Dalvi and James Glass ...	4638
<i>SenseBERT: Driving Some Sense into BERT</i> Yoav Levine, Barak Lenz, Or Dagan, Ori Ram, Dan Padnos, Or Sharir, Shai Shalev-Shwartz, Amnon Shashua and Yoav Shoham .....	4656
<i>ASSET: A Dataset for Tuning and Evaluation of Sentence Simplification Models with Multiple Rewriting Transformations</i> Fernando Alva-Manchego, Louis Martin, Antoine Bordes, Carolina Scarton, Benoît Sagot and Lucia Specia .....	4668
<i>Fatality Killed the Cat or: BabelPic, a Multimodal Dataset for Non-Concrete Concepts</i> Agostina Calabrese, Michele Bevilacqua and Roberto Navigli .....	4680
<i>Modeling Label Semantics for Predicting Emotional Reactions</i> Radhika Gaonkar, Heeyoung Kwon, Mohaddeseh Bastan, Niranjan Balasubramanian and Nathanael Chambers .....	4687

<i>CraftAssist Instruction Parsing: Semantic Parsing for a Voxel-World Assistant</i> Kavya Srinet, Yacine Jernite, Jonathan Gray and arthur szlam .....	4693
<i>Don't Say That! Making Inconsistent Dialogue Unlikely with Unlikelihood Training</i> Margaret Li, Stephen Roller, Ilya Kulikov, Sean Welleck, Y-Lan Boureau, Kyunghyun Cho and Jason Weston .....	4715
<i>How does BERT's attention change when you fine-tune? An analysis methodology and a case study in negation scope</i> Yiyun Zhao and Steven Bethard .....	4729
<i>Influence Paths for Characterizing Subject-Verb Number Agreement in LSTM Language Models</i> Kaiji Lu, Piotr Mardziel, Klas Leino, Matt Fredrikson and Anupam Datta .....	4748
<i>Interpreting Pretrained Contextualized Representations via Reductions to Static Embeddings</i> Rishi Bommasani, Kelly Davis and Claire Cardie .....	4758
<i>Learning to Deceive with Attention-Based Explanations</i> Danish Pruthi, Mansi Gupta, Bhuwan Dhingra, Graham Neubig and Zachary C. Lipton .....	4782
<i>On the Spontaneous Emergence of Discrete and Compositional Signals</i> Nur Geffen Lan, Emmanuel Chemla and Shane Steinert-Threlkeld .....	4794
<i>Spying on Your Neighbors: Fine-grained Probing of Contextual Embeddings for Information about Surrounding Words</i> Josef Klafka and Allyson Ettinger .....	4801
<i>Dense-Caption Matching and Frame-Selection Gating for Temporal Localization in VideoQA</i> Hyoungun Kim, Zineng Tang and Mohit Bansal .....	4812
<i>Shaping Visual Representations with Language for Few-Shot Classification</i> Jesse Mu, Percy Liang and Noah Goodman .....	4823
<i>Discrete Latent Variable Representations for Low-Resource Text Classification</i> Shuning Jin, Sam Wiseman, Karl Stratos and Karen Livescu .....	4831
<i>Learning Constraints for Structured Prediction Using Rectifier Networks</i> Xingyuan Pan, Maitrey Mehta and Vivek Srikumar .....	4843
<i>Pretraining with Contrastive Sentence Objectives Improves Discourse Performance of Language Models</i> Dan Iter, Kelvin Guu, Larry Lansing and Dan Jurafsky .....	4859
<i>A Recipe for Creating Multimodal Aligned Datasets for Sequential Tasks</i> Angela Lin, Sudha Rao, Asli Celikyilmaz, Elnaz Nouri, Chris Brockett, Debadepta Dey and Bill Dolan .....	4871
<i>Adversarial NLI: A New Benchmark for Natural Language Understanding</i> Yixin Nie, Adina Williams, Emily Dinan, Mohit Bansal, Jason Weston and Douwe Kiela .....	4885
<i>Beyond Accuracy: Behavioral Testing of NLP Models with CheckList</i> Marco Tulio Ribeiro, Tongshuang Wu, Carlos Guestrin and Sameer Singh .....	4902
<i>Code and Named Entity Recognition in StackOverflow</i> Jeniya Tabassum, Mounica Maddela, Wei Xu and Alan Ritter .....	4913

<i>Dialogue-Based Relation Extraction</i>	
Dian Yu, Kai Sun, Claire Cardie and Dong Yu .....	4927
<i>Facet-Aware Evaluation for Extractive Summarization</i>	
Yuning Mao, Liyuan Liu, Qi Zhu, Xiang Ren and Jiawei Han .....	4941
<i>More Diverse Dialogue Datasets via Diversity-Informed Data Collection</i>	
Katherine Stasaski, Grace Hui Yang and Marti A. Hearst .....	4958
<i>S2ORC: The Semantic Scholar Open Research Corpus</i>	
Kyle Lo, Lucy Lu Wang, Mark Neumann, Rodney Kinney and Daniel Weld .....	4969
<i>Tangled up in BLEU: Reevaluating the Evaluation of Automatic Machine Translation Evaluation Metrics</i>	
Nitika Mathur, Timothy Baldwin and Trevor Cohn .....	4984
<i>A Transformer-based Approach for Source Code Summarization</i>	
Wasi Ahmad, Saikat Chakraborty, Baishakhi Ray and Kai-Wei Chang .....	4998
<i>Asking and Answering Questions to Evaluate the Factual Consistency of Summaries</i>	
Alex Wang, Kyunghyun Cho and Mike Lewis .....	5008
<i>Discourse-Aware Neural Extractive Text Summarization</i>	
Jiacheng Xu, Zhe Gan, Yu Cheng and Jingjing Liu .....	5021
<i>Discrete Optimization for Unsupervised Sentence Summarization with Word-Level Extraction</i>	
Raphael Schumann, Lili Mou, Yao Lu, Olga Vechtomova and Katja Markert .....	5032
<i>Exploring Content Selection in Summarization of Novel Chapters</i>	
Faisal Ladhak, Bryan Li, Yaser Al-Onaizan and Kathleen McKeown .....	5043
<i>FEQA: A Question Answering Evaluation Framework for Faithfulness Assessment in Abstractive Summarization</i>	
Esin Durmus, He He and Mona Diab .....	5055
<i>Fact-based Content Weighting for Evaluating Abstractive Summarisation</i>	
Xinnuo Xu, Ondřej Dušek, Jingyi Li, Verena Rieser and Ioannis Konstas .....	5071
<i>Hooks in the Headline: Learning to Generate Headlines with Controlled Styles</i>	
Di Jin, Zhijing Jin, Joey Tianyi Zhou, Lisa Orii and Peter Szolovits .....	5082
<i>Knowledge Graph-Augmented Abstractive Summarization with Semantic-Driven Cloze Reward</i>	
Luyang Huang, Lingfei Wu and Lu Wang .....	5094
<i>Optimizing the Factual Correctness of a Summary: A Study of Summarizing Radiology Reports</i>	
Yuhao Zhang, Derek Merck, Emily Tsai, Christopher D. Manning and Curtis Langlotz .....	5108
<i>Storytelling with Dialogue: A Critical Role Dungeons and Dragons Dataset</i>	
Revanth Rameshkumar and Peter Bailey .....	5121
<i>The Summary Loop: Learning to Write Abstractive Summaries Without Examples</i>	
Philippe Laban, Andrew Hsi, John Canny and Marti A. Hearst .....	5135
<i>Unsupervised Opinion Summarization as Copycat-Review Generation</i>	
Arthur Bražinskas, Mirella Lapata and Ivan Titov .....	5151

<i>(Re)construing Meaning in NLP</i>	
Sean Trott, Tiago Timponi Torrent, Nancy Chang and Nathan Schneider .....	5170
<i>Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data</i>	
Emily M. Bender and Alexander Koller .....	5185
<i>Examining Citations of Natural Language Processing Literature</i>	
Saif M. Mohammad .....	5199
<i>How Can We Accelerate Progress Towards Human-like Linguistic Generalization?</i>	
Tal Linzen .....	5210
<i>How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence</i>	
Haoxi Zhong, Chaojun Xiao, Cunchao Tu, Tianyang Zhang, Zhiyuan Liu and Maosong Sun .	5218
<i>Intermediate-Task Transfer Learning with Pretrained Language Models: When and Why Does It Work?</i>	
Yada Pruksachatkun, Jason Phang, Haokun Liu, Phu Mon Htut, Xiaoyi Zhang, Richard Yuanzhe Pang, Clara Vania, Katharina Kann and Samuel R. Bowman .....	5231
<i>Predictive Biases in Natural Language Processing Models: A Conceptual Framework and Overview</i>	
Deven Santosh Shah, H. Andrew Schwartz and Dirk Hovy .....	5248
<i>What Does BERT with Vision Look At?</i>	
Liunian Harold Li, Mark Yatskar, Da Yin, Cho-Jui Hsieh and Kai-Wei Chang .....	5265
<i>Balancing Objectives in Counseling Conversations: Advancing Forwards or Looking Backwards</i>	
Justine Zhang and Cristian Danescu-Niculescu-Mizil .....	5276
<i>Detecting Perceived Emotions in Hurricane Disasters</i>	
Shrey Desai, Cornelia Caragea and Junyi Jessy Li .....	5290
<i>Hierarchical Modeling for User Personality Prediction: The Role of Message-Level Attention</i>	
Veronica Lynn, Niranjan Balasubramanian and H. Andrew Schwartz .....	5306
<i>Measuring Forecasting Skill from Text</i>	
Shi Zong, Alan Ritter and Eduard Hovy .....	5317
<i>Text and Causal Inference: A Review of Using Text to Remove Confounding from Causal Estimates</i>	
Katherine Keith, David Jensen and Brendan O'Connor .....	5332
<i>Text-Based Ideal Points</i>	
Keyon Vafa, Suresh Naidu and David Blei .....	5345
<i>Understanding the Language of Political Agreement and Disagreement in Legislative Texts</i>	
Maryam Davoodi, Eric Waltenburg and Dan Goldwasser .....	5358
<i>Would you Rather? A New Benchmark for Learning Machine Alignment with Cultural Values and Social Preferences</i>	
Yi Tay, Donovan Ong, Jie Fu, Alvin Chan, Nancy Chen, Anh Tuan Luu and Chris Pal .....	5369
<i>Discourse as a Function of Event: Profiling Discourse Structure in News Articles around the Main Event</i>	
Prafulla Kumar Choubey, Aaron Lee, Ruihong Huang and Lu Wang .....	5374
<i>Harnessing the linguistic signal to predict scalar inferences</i>	
Sebastian Schuster, Yuxing Chen and Judith Degen .....	5387

<i>Implicit Discourse Relation Classification: We Need to Talk about Evaluation</i>	
Najoung Kim, Song Feng, Chulaka Gunasekara and Luis Lastras .....	5404
<i>PeTra: A Sparsely Supervised Memory Model for People Tracking</i>	
Shubham Toshniwal, Allyson Ettinger, Kevin Gimpel and Karen Livescu .....	5415
<i>ZPR2: Joint Zero Pronoun Recovery and Resolution using Multi-Task Learning and BERT</i>	
Lin Feng Song, Kun Xu, Yue Zhang, Jianshu Chen and Dong Yu .....	5429
<i>Contextualizing Hate Speech Classifiers with Post-hoc Explanation</i>	
Brendan Kennedy, Xisen Jin, Aida Mostafazadeh Davani, Morteza Dehghani and Xiang Ren .....	5435
<i>Double-Hard Debias: Tailoring Word Embeddings for Gender Bias Mitigation</i>	
Tianlu Wang, Xi Victoria Lin, Nazneen Fatema Rajani, Bryan McCann, Vicente Ordonez and Caiming Xiong .....	5443
<i>Language (Technology) is Power: A Critical Survey of "Bias" in NLP</i>	
Su Lin Blodgett, Solon Barocas, Hal Daumé III and Hanna Wallach .....	5454
<i>Social Bias Frames: Reasoning about Social and Power Implications of Language</i>	
Maarten Sap, Saadia Gabriel, Lianhui Qin, Dan Jurafsky, Noah A. Smith and Yejin Choi .....	5477
<i>Social Biases in NLP Models as Barriers for Persons with Disabilities</i>	
Ben Hutchinson, Vinodkumar Prabhakaran, Emily Denton, Kellie Webster, Yu Zhong and Stephen Denuyl .....	5491
<i>Towards Debiasing Sentence Representations</i>	
Paul Pu Liang, Irene Mengze Li, Emily Zheng, Yao Chong Lim, Ruslan Salakhutdinov and Louis-Philippe Morency .....	5502
<i>A Re-evaluation of Knowledge Graph Completion Methods</i>	
Zhiqing Sun, Shikhar Vashishth, Soumya Sanyal, Partha Talukdar and Yiming Yang .....	5516
<i>Cross-Linguistic Syntactic Evaluation of Word Prediction Models</i>	
Aaron Mueller, Garrett Nicolai, Panayiota Petrou-Zeniou, Natalia Talmina and Tal Linzen .....	5523
<i>Evaluating Explainable AI: Which Algorithmic Explanations Help Users Predict Model Behavior?</i>	
Peter Hase and Mohit Bansal .....	5540
<i>Explaining Black Box Predictions and Unveiling Data Artifacts through Influence Functions</i>	
Xiaochuang Han, Byron C. Wallace and Yulia Tsvetkov .....	5553
<i>Finding Universal Grammatical Relations in Multilingual BERT</i>	
Ethan A. Chi, John Hewitt and Christopher D. Manning .....	5564
<i>Generating Hierarchical Explanations on Text Classification via Feature Interaction Detection</i>	
Hanjie Chen, Guangtao Zheng and Yangfeng Ji .....	5578
<i>Obtaining Faithful Interpretations from Compositional Neural Networks</i>	
Sanjay Subramanian, Ben Bogin, Nitish Gupta, Tomer Wolfson, Sameer Singh, Jonathan Berant and Matt Gardner .....	5594
<i>Rationalizing Text Matching: Learning Sparse Alignments via Optimal Transport</i>	
Kyle Swanson, Lili Yu and Tao Lei .....	5609



<i>Benefits of Intermediate Annotations in Reading Comprehension</i>	
Dheeru Dua, Sameer Singh and Matt Gardner .....	5627
<i>Crossing Variational Autoencoders for Answer Retrieval</i>	
Wenhao Yu, Lingfei Wu, Qingkai Zeng, Shu Tao, Yu Deng and Meng Jiang .....	5635
<i>Logic-Guided Data Augmentation and Regularization for Consistent Question Answering</i>	
Akari Asai and Hannaneh Hajishirzi .....	5642
<i>On the Importance of Diversity in Question Generation for QA</i>	
Md Arifat Sultan, Shubham Chandel, Ramón Fernandez Astudillo and Vittorio Castelli .....	5651
<i>Probabilistic Assumptions Matter: Improved Models for Distantly-Supervised Document-Level Question Answering</i>	
Hao Cheng, Ming-Wei Chang, Kenton Lee and Kristina Toutanova .....	5657
<i>SCDE: Sentence Cloze Dataset with High Quality Distractors From Examinations</i>	
Xiang Kong, Varun Gangal and Eduard Hovy .....	5668
<i>Selective Question Answering under Domain Shift</i>	
Amita Kamath, Robin Jia and Percy Liang .....	5684
<i>The Cascade Transformer: an Application for Efficient Answer Sentence Selection</i>	
Luca Soldaini and Alessandro Moschitti .....	5697
<i>Transformers to Learn Hierarchical Contexts in Multiparty Dialogue for Span-based Question Answering</i>	
Changmao Li and Jinho D. Choi .....	5709
<i>Not All Claims are Created Equal: Choosing the Right Statistical Approach to Assess Hypotheses</i>	
Erfan Sadeqi Azer, Daniel Khashabi, Ashish Sabharwal and Dan Roth .....	5715
<i>STARC: Structured Annotations for Reading Comprehension</i>	
Yevgeni Berzak, Jonathan Malmaud and Roger Levy .....	5726
<i>WinoWhy: A Deep Diagnosis of Essential Commonsense Knowledge for Answering Winograd Schema Challenge</i>	
Hongming Zhang, Xinran Zhao and Yangqiu Song .....	5736
<i>Agreement Prediction of Arguments in Cyber Argumentation for Detecting Stance Polarity and Intensity</i>	
Joseph Sirrianni, Xiaoqing Liu and Douglas Adams .....	5746
<i>Cross-Lingual Unsupervised Sentiment Classification with Multi-View Transfer Learning</i>	
Hongliang Fei and Ping Li .....	5759
<i>Efficient Pairwise Annotation of Argument Quality</i>	
Lukas Gienapp, Benno Stein, Matthias Hagen and Martin Potthast .....	5772
<i>Entity-Aware Dependency-Based Deep Graph Attention Network for Comparative Preference Classification</i>	
Nianzu Ma, Sahisnu Mazumder, Hao Wang and Bing Liu .....	5782
<i>OpinionDigest: A Simple Framework for Opinion Summarization</i>	
Yoshihiko Suhara, Xiaolan Wang, Stefanos Angelidis and Wang-Chiew Tan .....	5789



<i>A Comprehensive Analysis of Preprocessing for Word Representation Learning in Affective Tasks</i> Nastaran Babanejad, Ameeta Agrawal, Aijun An and Manos Papagelis .....	5799
<i>Diverse and Informative Dialogue Generation with Context-Specific Commonsense Knowledge Awareness</i> Sixing Wu, Ying Li, Dawei Zhang, Yang Zhou and Zhonghai Wu .....	5811
<i>Generate, Delete and Rewrite: A Three-Stage Framework for Improving Persona Consistency of Dialogue Generation</i> Haoyu Song, Yan Wang, Wei-Nan Zhang, Xiaojiang Liu and Ting Liu .....	5821
<i>Learning to Customize Model Structures for Few-shot Dialogue Generation Tasks</i> YIPING SONG, Zequn Liu, Wei Bi, Rui Yan and Ming Zhang .....	5832
<i>Video-Grounded Dialogues with Pretrained Generation Language Models</i> Hung Le and Steven C.H. Hoi .....	5842
<i>A Unified MRC Framework for Named Entity Recognition</i> Xiaoya Li, Jingrong Feng, Yuxian Meng, Qinghong Han, Fei Wu and Jiwei Li .....	5849
<i>An Effective Transition-based Model for Discontinuous NER</i> Xiang Dai, Sarvnaz Karimi, Ben Hachey and Cecile Paris .....	5860
<i>IMoJIE: Iterative Memory-Based Joint Open Information Extraction</i> Keshav Kolluru, Samarth Aggarwal, Vipul Rathore, and Soumen Chakrabarti .....	5871
<i>Improving Event Detection via Open-domain Trigger Knowledge</i> Meihan Tong, Bin Xu, Shuai Wang, Yixin Cao, Lei Hou, Juanzi Li and Jun Xie .....	5887
<i>Improving Low-Resource Named Entity Recognition using Joint Sentence and Token Labeling</i> Canasai Kruengkrai, Thien Hai Nguyen, Sharifah Mahani Aljunied and Lidong Bing .....	5898
<i>Multi-Cell Compositional LSTM for NER Domain Adaptation</i> Chen Jia and Yue Zhang .....	5906
<i>Pyramid: A Layered Model for Nested Named Entity Recognition</i> Jue WANG, Lidan Shou, Ke Chen and Gang Chen .....	5918
<i>ReInceptionE: Relation-Aware Inception Network with Joint Local-Global Structural Information for Knowledge Graph Embedding</i> Zhiwen Xie, Guangyou Zhou, Jin Liu and Jimmy Xiangji Huang .....	5929
<i>Relabel the Noise: Joint Extraction of Entities and Relations via Cooperative Multiagents</i> Daoyuan Chen, Yaliang Li, Kai Lei and Ying Shen .....	5940
<i>Simplify the Usage of Lexicon in Chinese NER</i> Ruotian Ma, Minlong Peng, Qi Zhang, Zhongyu Wei and Xuanjing Huang .....	5951
<i>AdvAug: Robust Adversarial Augmentation for Neural Machine Translation</i> Yong Cheng, Lu Jiang, Wolfgang Macherey and Jacob Eisenstein .....	5961
<i>Contextual Neural Machine Translation Improves Translation of Cataphoric Pronouns</i> KayYen Wong, Sameen Maruf and Gholamreza Haffari .....	5971
<i>Improving Neural Machine Translation with Soft Template Prediction</i> Jian Yang, Shuming Ma, Dongdong Zhang, Zhoujun Li and Ming Zhou .....	5979

<i>Tagged Back-translation Revisited: Why Does It Really Work?</i>	
Benjamin Marie, Raphael Rubino and Atsushi Fujita .....	5990
<i>Worse WER, but Better BLEU? Leveraging Word Embedding as Intermediate in Multitask End-to-End Speech Translation</i>	
Shun-Po Chuang, Tzu-Wei Sung, Alexander H. Liu and Hung-yi Lee .....	5998
<i>Neural-DINF: A Neural Network based Framework for Measuring Document Influence</i>	
Jie Tan, Changlin Yang, Ying Li, Siliang Tang, Chen Huang and Yueting Zhuang .....	6004
<i>Paraphrase Generation by Learning How to Edit from Samples</i>	
Amirhossein Kazemnejad, Mohammadreza Salehi and Mahdieh Soleymani Baghshah .....	6010
<i>Emerging Cross-lingual Structure in Pretrained Language Models</i>	
Alexis Conneau, Shijie Wu, Haoran Li, Luke Zettlemoyer and Veselin Stoyanov .....	6022
<i>FastBERT: a Self-distilling BERT with Adaptive Inference Time</i>	
Weijie Liu, Peng Zhou, Zhiruo Wang, Zhe Zhao, Haotang Deng and QI JU .....	6035
<i>Incorporating External Knowledge through Pre-training for Natural Language to Code Generation</i>	
Frank F. Xu, Zhengbao Jiang, Pengcheng Yin, Bogdan Vasilescu and Graham Neubig .....	6045
<i>LogicalFactChecker: Leveraging Logical Operations for Fact Checking with Graph Module Network</i>	
Wanjun Zhong, Duyu Tang, Zhangyin Feng, Nan Duan, Ming Zhou, Ming Gong, Linjun Shou, Daxin Jiang, Jiahai Wang and Jian Yin .....	6053
<i>Word-level Textual Adversarial Attacking as Combinatorial Optimization</i>	
Yuan Zang, Fanchao Qi, Chenghao Yang, Zhiyuan Liu, Meng Zhang, Qun Liu and Maosong Sun	6066
<i>Benchmarking Multimodal Regex Synthesis with Complex Structures</i>	
Xi Ye, Qiaochu Chen, Isil Dillig and Greg Durrett .....	6081
<i>Curriculum Learning for Natural Language Understanding</i>	
Benfeng Xu, Licheng Zhang, Zhendong Mao, Quan Wang, Hongtao Xie and Yongdong Zhang	6095
<i>Do Neural Models Learn Systematicity of Monotonicity Inference in Natural Language?</i>	
Hitomi Yanaka, Koji Mineshima, Daisuke Bekki and Kentaro Inui .....	6105
<i>Evidence-Aware Inferential Text Generation with Vector Quantised Variational AutoEncoder</i>	
Daya Guo, Duyu Tang, Nan Duan, Jian Yin, Daxin Jiang and Ming Zhou .....	6118
<i>How to Ask Good Questions? Try to Leverage Paraphrases</i>	
Xin Jia, Wenjie Zhou, Xu SUN and Yunfang Wu .....	6130
<i>NeuInfer: Knowledge Inference on N-ary Facts</i>	
Saiping Guan, Xiaolong Jin, Jiafeng Guo, Yuanzhuo Wang and Xueqi Cheng .....	6141
<i>Neural Graph Matching Networks for Chinese Short Text Matching</i>	
Lu Chen, Yanbin Zhao, Boer Lyu, Lesheng Jin, Zhi Chen, Su Zhu and Kai Yu .....	6152
<i>Neural Mixed Counting Models for Dispersed Topic Discovery</i>	
Jiemin Wu, Yanghui Rao, Zusheng Zhang, Haoran Xie, Qing Li, Fu Lee Wang and Ziyi Chen	6159

<i>Reasoning Over Semantic-Level Graph for Fact Checking</i>	
Wanjun Zhong, Jingjing Xu, Duyu Tang, Zenan Xu, Nan Duan, Ming Zhou, Jiahai Wang and Jian Yin .....	6170
<i>Automatic Generation of Citation Texts in Scholarly Papers: A Pilot Study</i>	
Xinyu Xing, Xiaosheng Fan and Xiaojun Wan .....	6181
<i>Composing Elementary Discourse Units in Abstractive Summarization</i>	
Zhenwen Li, Wenhao Wu and Sujian Li .....	6191
<i>Extractive Summarization as Text Matching</i>	
Ming Zhong, Pengfei Liu, Yiran Chen, Danqing Wang, Xipeng Qiu and Xuanjing Huang ....	6197
<i>Heterogeneous Graph Neural Networks for Extractive Document Summarization</i>	
Danqing Wang, Pengfei Liu, Yining Zheng, Xipeng Qiu and Xuanjing Huang .....	6209
<i>Jointly Learning to Align and Summarize for Neural Cross-Lingual Summarization</i>	
Yue Cao, Hui Liu and Xiaojun Wan .....	6220
<i>Leveraging Graph to Improve Abstractive Multi-Document Summarization</i>	
Wei Li, Xinyan Xiao, Jiachen Liu, Hua Wu, Haifeng Wang and Junping Du .....	6232
<i>Multi-Granularity Interaction Network for Extractive and Abstractive Multi-Document Summarization</i>	
Hanqi Jin, Tianming Wang and Xiaojun Wan .....	6244
<i>Tetra-Tagging: Word-Synchronous Parsing with Linear-Time Inference</i>	
Nikita Kitaev and Dan Klein .....	6255
<i>Are we Estimating or Guesstimating Translation Quality?</i>	
Shuo Sun, Francisco Guzmán and Lucia Specia .....	6262
<i>Language (Re)modelling: Towards Embodied Language Understanding</i>	
Ronen Tamari, Chen Shani, Tom Hope, Miriam R L Petruck, Omri Abend and Dafna Shahaf .	6268
<i>The State and Fate of Linguistic Diversity and Inclusion in the NLP World</i>	
Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali and Monojit Choudhury .....	6282
<i>The Unstoppable Rise of Computational Linguistics in Deep Learning</i>	
James Henderson .....	6294
<i>To Boldly Query What No One Has Annotated Before? The Frontiers of Corpus Querying</i>	
Markus Gärtner and Kerstin Jung .....	6307
<i>A Contextual Hierarchical Attention Network with Adaptive Objective for Dialogue State Tracking</i>	
Yong Shan, Zekang Li, Jinchao Zhang, Fandong Meng, Yang Feng, Cheng Niu and Jie Zhou .	6322
<i>Data Manipulation: Towards Effective Instance Learning for Neural Dialogue Generation via Learning to Augment and Reweight</i>	
Hengyi Cai, Hongshen Chen, Yonghao Song, Cheng Zhang, Xiaofang Zhao and Dawei Yin . .	6334
<i>Dynamic Fusion Network for Multi-Domain End-to-end Task-Oriented Dialog</i>	
Libo Qin, Xiao Xu, Wanxiang Che, Yue Zhang and Ting Liu .....	6344
<i>Learning Efficient Dialogue Policy from Demonstrations through Shaping</i>	
Huimin Wang, Baolin Peng and Kam-Fai Wong .....	6355

<i>SAS: Dialogue State Tracking via Slot Attention and Slot Information Sharing</i>	
Jiaying Hu, Yan Yang, Chencai Chen, liang he and Zhou Yu .....	6366
<i>Speaker Sensitive Response Evaluation Model</i>	
JinYeong Bak and Alice Oh .....	6376
<i>A Top-down Neural Architecture towards Text-level Parsing of Discourse Rhetorical Structure</i>	
Longyin Zhang, Yuqing Xing, Fang Kong, Peifeng Li and Guodong Zhou .....	6386
<i>Amalgamation of protein sequence, structure and textual information for improving protein-protein interaction identification</i>	
pratik Dutta and Sriparna Saha .....	6396
<i>Bipartite Flat-Graph Network for Nested Named Entity Recognition</i>	
Ying Luo and Hai Zhao .....	6408
<i>Connecting Embeddings for Knowledge Graph Entity Typing</i>	
Yu Zhao, anxiang zhang, Ruobing Xie, Kang Liu and Xiaojie WANG .....	6419
<i>Continual Relation Learning via Episodic Memory Activation and Reconsolidation</i>	
Xu Han, Yi Dai, Tianyu Gao, Yankai Lin, Zhiyuan Liu, Peng Li, Maosong Sun and Jie Zhou ..	6429
<i>Handling Rare Entities for Neural Sequence Labeling</i>	
Yangming Li, Han Li, Kaisheng Yao and Xiaolong Li .....	6441
<i>Instance-Based Learning of Span Representations: A Case Study through Named Entity Recognition</i>	
Hiroki Ouchi, Jun Suzuki, Sosuke Kobayashi, Sho Yokoi, Tatsuki Kuribayashi, Ryuto Konno and Kentaro Inui .....	6452
<i>MIE: A Medical Information Extractor towards Medical Dialogues</i>	
Yuanzhe Zhang, Zhongtao Jiang, Tao Zhang, Shiwang Liu, Jiarun Cao, Kang Liu, Shengping Liu and Jun Zhao .....	6460
<i>Named Entity Recognition as Dependency Parsing</i>	
Juntao Yu, Bernd Bohnet and Massimo Poesio .....	6470
<i>Neighborhood Matching Network for Entity Alignment</i>	
Yuting Wu, Xiao Liu, Yansong Feng, Zheng Wang and Dongyan Zhao .....	6477
<i>Relation Extraction with Explanation</i>	
Hamed Shahbazi, Xiaoli Fern, Reza Ghaeini and Prasad Tadepalli .....	6488
<i>Representation Learning for Information Extraction from Form-like Documents</i>	
Bodhisattwa Prasad Majumder, Navneet Potti, Sandeep Tata, James Bradley Wendt, Qi Zhao and Marc Najork .....	6495
<i>Single-/Multi-Source Cross-Lingual NER via Teacher-Student Learning on Unlabeled Data in Target Language</i>	
Qianhui Wu, Zijia Lin, Börje Karlsson, Jian-Guang LOU and Biqing Huang .....	6505
<i>Synchronous Double-channel Recurrent Network for Aspect-Opinion Pair Extraction</i>	
Shaowei Chen, Jie Liu, Yu Wang, Wenzheng Zhang and Ziming Chi .....	6515
<i>Cross-modal Coherence Modeling for Caption Generation</i>	
Malihe Alikhani, Piyush Sharma, Shengjie Li, Radu Soricut and Matthew Stone .....	6525

<i>Knowledge Supports Visual Language Grounding: A Case Study on Colour Terms</i> Simeon Schüz and Sina Zarriß .....	6536
<i>Span-based Localizing Network for Natural Language Video Localization</i> Hao Zhang, Aixin Sun, Wei Jing and Joey Tianyi Zhou .....	6543
<i>Words Aren't Enough, Their Order Matters: On the Robustness of Grounding Visual Referring Expressions</i> Arjun Akula, Spandana Gella, Yaser Al-Onaizan, Song-Chun Zhu and Siva Reddy .....	6555
<i>A Mixture of <math>h - 1</math> Heads is Better than <math>h</math> Heads</i> Hao Peng, Roy Schwartz, Dianqi Li and Noah A. Smith .....	6566
<i>Dependency Graph Enhanced Dual-transformer Structure for Aspect-based Sentiment Classification</i> Hao Tang, Donghong Ji, Chenliang Li and Qiji Zhou .....	6578
<i>Differentiable Window for Dynamic Local Attention</i> Thanh-Tung Nguyen, Xuan-Phi Nguyen, Shafiq Joty and Xiaoli Li .....	6589
<i>Evaluating and Enhancing the Robustness of Neural Network-based Dependency Parsing Models with Adversarial Examples</i> Xiaoqing Zheng, Jiehang Zeng, Yi Zhou, Cho-Jui Hsieh, Minhao Cheng and Xuanjing Huang	6600
<i>Exploiting Syntactic Structure for Better Language Modeling: A Syntactic Distance Approach</i> Wenyu Du, Zhouhan Lin, Yikang Shen, Timothy J. O'Donnell, Yoshua Bengio and Yue Zhang	6611
<i>Learning Architectures from an Extended Search Space for Language Modeling</i> Yinqiao Li, Chi Hu, Yuhao Zhang, Nuo Xu, Yufan Jiang, Tong Xiao, Jingbo Zhu, Tongran Liu and changliang li .....	6629
<i>The Right Tool for the Job: Matching Model and Instance Complexities</i> Roy Schwartz, Gabriel Stanovsky, Swabha Swayamdipta, Jesse Dodge and Noah A. Smith ..	6640
<i>Bootstrapping Techniques for Polysynthetic Morphological Analysis</i> William Lane and Steven Bird .....	6652
<i>Coupling Distant Annotation and Adversarial Training for Cross-Domain Chinese Word Segmentation</i> Ning Ding, Dingkun Long, Guangwei Xu, Muhua Zhu, Pengjun Xie, Xiaobin Wang and Haitao Zheng .....	6662
<i>Modeling Morphological Typology for Unsupervised Learning of Language Morphology</i> Hongzhi Xu, Jordan Kodner, Mitchell Marcus and Charles Yang .....	6672
<i>Predicting Declension Class from Form and Meaning</i> Adina Williams, Tiago Pimentel, Hagen Blix, Arya D. McCarthy, Eleanor Chodroff and Ryan Cotterell .....	6682
<i>Unsupervised Morphological Paradigm Completion</i> Huiming Jin, Liwei Cai, Yihui Peng, Chen Xia, Arya McCarthy and Katharina Kann .....	6696
<i>Document Modeling with Graph Attention Networks for Multi-grained Machine Reading Comprehension</i> Bo Zheng, Haoyang Wen, Yaobo Liang, Nan Duan, Wanxiang Che, Daxin Jiang, Ming Zhou and Ting Liu .....	6708
<i>Harvesting and Refining Question-Answer Pairs for Unsupervised QA</i> Zhongli Li, Wenhui Wang, Li Dong, Furu Wei and Ke Xu .....	6719

<i>Low-Resource Generation of Multi-hop Reasoning Questions</i>	
Jianxing Yu, Wei Liu, Shuang Qiu, Qinliang Su, Kai Wang, Xiaojun Quan and Jian Yin . . . . .	6729
<i>R4C: A Benchmark for Evaluating RC Systems to Get the Right Answer for the Right Reason</i>	
Naoya Inoue, Pontus Stenetorp and Kentaro Inui . . . . .	6740
<i>Recurrent Chunking Mechanisms for Long-Text Machine Reading Comprehension</i>	
Hongyu Gong, Yelong Shen, Dian Yu, Jianshu Chen and Dong Yu . . . . .	6751
<i>RikiNet: Reading Wikipedia Pages for Natural Question Answering</i>	
Dayiheng Liu, Yeyun Gong, Jie Fu, Yu Yan, Jiusheng Chen, Daxin Jiang, Jiancheng Lv and Nan Duan . . . . .	6762
<i>Parsing into Variable-in-situ Logico-Semantic Graphs</i>	
Yufei Chen and Weiwei Sun . . . . .	6772
<i>Semantic Parsing for English as a Second Language</i>	
Yuanyuan Zhao, Weiwei Sun, Junjie Cao and Xiaojun Wan . . . . .	6783
<i>Semi-Supervised Semantic Dependency Parsing Using CRF Autoencoders</i>	
Zixia Jia, Youmi Ma, Jiong Cai and Kewei Tu . . . . .	6795
<i>Unsupervised Dual Paraphrasing for Two-stage Semantic Parsing</i>	
Ruisheng Cao, Su Zhu, Chenyu Yang, Chen Liu, Rao Ma, Yanbin Zhao, Lu Chen and Kai Yu . . . . .	6806
<i>DRTS Parsing with Structure-Aware Encoding and Decoding</i>	
Qiankun Fu, Yue Zhang, Jiangming Liu and Meishan Zhang . . . . .	6818
<i>A Two-Stage Masked LM Method for Term Set Expansion</i>	
Guy Kushilevitz, Shaul Markovitch and Yoav Goldberg . . . . .	6829
<i>FLAT: Chinese NER Using Flat-Lattice Transformer</i>	
Xiaonan Li, Hang Yan, Xipeng Qiu and Xuanjing Huang . . . . .	6836
<i>Improving Entity Linking through Semantic Reinforced Entity Embeddings</i>	
Feng Hou, Ruili Wang, Jun He and Yi Zhou . . . . .	6843
<i>Document Translation vs. Query Translation for Cross-Lingual Information Retrieval in the Medical Domain</i>	
Shadi Saleh and Pavel Pecina . . . . .	6849
<i>Learning Robust Models for e-Commerce Product Search</i>	
Thanh Nguyen, Nikhil Rao and Karthik Subbian . . . . .	6861
<i>Generalized Entropy Regularization or: There's Nothing Special about Label Smoothing</i>	
Clara Meister, Elizabeth Salesky and Ryan Cotterell . . . . .	6870
<i>Highway Transformer: Self-Gating Enhanced Self-Attentive Networks</i>	
Yekun Chai, Shuo Jin and Xinwen Hou . . . . .	6887
<i>Low-Dimensional Hyperbolic Knowledge Graph Embeddings</i>	
Ines Chami, Adva Wolf, Da-Cheng Juan, Frederic Sala, Sujith Ravi and Christopher Ré . . . . .	6901
<i>Classification-Based Self-Learning for Weakly Supervised Bilingual Lexicon Induction</i>	
Mladen Karan, Ivan Vulić, Anna Korhonen and Goran Glavaš . . . . .	6915

<i>Gender in Danger? Evaluating Speech Translation Technology on the MuST-SHE Corpus</i>	
Luisa Bentivogli, Beatrice Savoldi, Matteo Negri, Mattia A. Di Gangi, Roldano Cattoni and Marco Turchi .....	6923
<i>Uncertainty-Aware Curriculum Learning for Neural Machine Translation</i>	
Yikai Zhou, Baosong Yang, Derek F. Wong, Yu Wan and Lidia S. Chao .....	6934
<i>Closing the Gap: Joint De-Identification and Concept Extraction in the Clinical Domain</i>	
Lukas Lange, Heike Adel and Jannik Strötgen .....	6945
<i>CorefQA: Coreference Resolution as Query-based Span Prediction</i>	
Wei Wu, Fei Wang, Arianna Yuan, Fei Wu and Jiwei Li .....	6953
<i>Estimating predictive uncertainty for rumour verification models</i>	
Elena Kochkina and Maria Liakata .....	6964
<i>From Zero to Hero: Human-In-The-Loop Entity Linking in Low Resource Domains</i>	
Jan-Christoph Klie, Richard Eckart de Castilho and Iryna Gurevych .....	6982
<i>Language to Network: Conditional Parameter Adaptation with Natural Language Descriptions</i>	
Tian Jin, Zhun Liu, Shengjia Yan, Alexandre Eichenberger and Louis-Philippe Morency .....	6994
<i>Controlled Crowdsourcing for High-Quality QA-SRL Annotation</i>	
Paul Roit, Ayal Klein, Daniela Stepanov, Jonathan Mamou, Julian Michael, Gabriel Stanovsky, Luke Zettlemoyer and Ido Dagan .....	7008
<i>Cross-Lingual Semantic Role Labeling with High-Quality Translated Training Corpus</i>	
Hao Fei, Meishan Zhang and Donghong Ji .....	7014
<i>Sentence Meta-Embeddings for Unsupervised Semantic Textual Similarity</i>	
Nina Poerner, Ulli Waltinger and Hinrich Schütze .....	7027
<i>Transition-based Semantic Dependency Parsing with Pointer Networks</i>	
Daniel Fernández-González and Carlos Gómez-Rodríguez .....	7035
<i>tBERT: Topic Models and BERT Joining Forces for Semantic Similarity Detection</i>	
Nicole Peinelt, Dong Nguyen and Maria Liakata .....	7047
<i>Conditional Augmentation for Aspect Term Extraction via Masked Sequence-to-Sequence Generation</i>	
Kun Li, Chengbo Chen, Xiaojun Quan, Qing Ling and Yan Song .....	7056
<i>Exploiting Personal Characteristics of Debaters for Predicting Persuasiveness</i>	
Khalid Al Khatib, Michael Völske, Shahbaz Syed, Nikolay Kolyada and Benno Stein .....	7067
<i>Out of the Echo Chamber: Detecting Countering Debate Speeches</i>	
Matan Orbach, Yonatan Bilu, Assaf Toledo, Dan Lahav, Michal Jacovi, Ranit Aharonov and Noam Slonim .....	7073
<i>Diversifying Dialogue Generation with Non-Conversational Text</i>	
Hui Su, Xiaoyu Shen, Sanqiang Zhao, Zhou Xiao, Pengwei Hu, randy zhong, Cheng Niu and Jie Zhou .....	7087
<i>KdConv: A Chinese Multi-domain Dialogue Dataset Towards Multi-turn Knowledge-driven Conversation</i>	
Hao Zhou, Chuji Zheng, Kaili Huang, Minlie Huang and Xiaoyan Zhu .....	7098



<i>Meta-Reinforced Multi-Domain State Generator for Dialogue Systems</i>	
Yi Huang, Junlan Feng, Min Hu, Xiaoting Wu, Xiaoyu Du and Shuo Ma .....	7109
<i>Modeling Long Context for Task-Oriented Dialogue State Generation</i>	
Jun Quan and Deyi Xiong .....	7119
<i>Multi-Domain Dialogue Acts and Response Co-Generation</i>	
Kai Wang, Junfeng Tian, Rui Wang, Xiaojun Quan and Jianxing Yu .....	7125
<i>Exploring Contextual Word-level Style Relevance for Unsupervised Style Transfer</i>	
Chulun Zhou, Liangyu Chen, Jiachen Liu, Xinyan Xiao, Jinsong Su, Sheng Guo and Hua Wu	7135
<i>Heterogeneous Graph Transformer for Graph-to-Sequence Learning</i>	
Shaowei Yao, Tianming Wang and Xiaojun Wan .....	7145
<i>Neural Data-to-Text Generation via Jointly Learning the Segmentation and Correspondence</i>	
Xiaoyu Shen, Ernie Chang, Hui Su, Cheng Niu and Dietrich Klakow .....	7155
<i>Aligned Dual Channel Graph Convolutional Network for Visual Question Answering</i>	
Qingbao Huang, Jielong Wei, Yi Cai, Changmeng Zheng, Junying Chen, Ho-fung Leung and Qing Li .....	7166
<i>Multimodal Neural Graph Memory Networks for Visual Question Answering</i>	
Mahmoud Khademi .....	7177
<i>Refer360°: A Referring Expression Recognition Dataset in 360° Images</i>	
Volkan Cirik, Taylor Berg-Kirkpatrick and Louis-Philippe Morency .....	7189
<i>CamemBERT: a Tasty French Language Model</i>	
Louis Martin, Benjamin Muller, Pedro Javier Ortiz Suárez, Yoann Dupont, Laurent Romary, Éric de la Clergerie, Djamel Seddah and Benoît Sagot .....	7203
<i>Effective Estimation of Deep Generative Language Models</i>	
Tom Pelsmaecker and Wilker Aziz .....	7220
<i>Null It Out: Guarding Protected Attributes by Iterative Nullspace Projection</i>	
Shauli Ravfogel, Yanai Elazar, Hila Gonen, Michael Twiton and Yoav Goldberg .....	7237
<i>2kenize: Tying Subword Sequences for Chinese Script Conversion</i>	
and Isabelle Augenstein .....	7257
<i>Predicting the Growth of Morphological Families from Social and Linguistic Factors</i>	
Valentin Hofmann, Janet Pierrehumbert and Hinrich Schütze .....	7273
<i>Semi-supervised Contextual Historical Text Normalization</i>	
Peter Makarov and Simon Clematide .....	7284
<i>ClarQ: A large-scale and diverse dataset for Clarification Question Generation</i>	
Vaibhav Kumar and Alan W Black .....	7296
<i>DoQA - Accessing Domain-Specific FAQs via Conversational QA</i>	
Jon Ander Campos, Arantxa Otegi, Aitor Soroa, Jan Deriu, Mark Cieliebak and Eneko Agirre	7302
<i>MLQA: Evaluating Cross-lingual Extractive Question Answering</i>	
Patrick Lewis, Barlas Oguz, Ruty Rinott, Sebastian Riedel and Holger Schwenk .....	7315



<i>Multi-source Meta Transfer for Low Resource Multiple-Choice Question Answering</i>	
Ming Yan, Hao Zhang, Di Jin and Joey Tianyi Zhou .....	7331
<i>Fine-grained Fact Verification with Kernel Graph Attention Network</i>	
Zhenghao Liu, Chenyan Xiong, Maosong Sun and Zhiyuan Liu .....	7342
<i>Generating Fact Checking Explanations</i>	
Pepa Atanasova, Jakob Grue Simonsen, Christina Lioma and Isabelle Augenstein .....	7352
<i>Premise Selection in Natural Language Mathematical Texts</i>	
Deborah Ferreira and André Freitas .....	7365
<i>A Call for More Rigor in Unsupervised Cross-lingual Learning</i>	
Mikel Artetxe, Sebastian Ruder, Dani Yogatama, Gorka Labaka and Eneko Agirre .....	7375
<i>A Tale of a Probe and a Parser</i>	
Rowan Hall Maudslay, Josef Valvoda, Tiago Pimentel, Adina Williams and Ryan Cotterell ..	7389
<i>From SPMRL to NMRL: What Did We Learn (and Unlearn) in a Decade of Parsing Morphologically-Rich Languages (MRLs)?</i>	
Reut Tsarfaty, Dan Baret, Stav Klein and Amit Seker .....	7396
<i>Speech Translation and the End-to-End Promise: Taking Stock of Where We Are</i>	
Matthias Sperber and Matthias Paulik .....	7409
<i>What Question Answering can Learn from Trivia Nerds</i>	
Jordan Boyd-Graber and Benjamin Börschinger .....	7422
<i>What are the Goals of Distributional Semantics?</i>	
Guy Emerson .....	7436
<i>Improving Image Captioning with Better Use of Caption</i>	
Zhan Shi, Xu Zhou, Xipeng Qiu and Xiaodan Zhu .....	7454
<i>Shape of Synth to Come: Why We Should Use Synthetic Data for English Surface Realization</i>	
Henry Elder, Robert Burke, Alexander O'Connor and Jennifer Foster .....	7465
<i>Toward Better Storylines with Sentence-Level Language Models</i>	
Daphne Ippolito, David Grangier, Douglas Eck and Chris Callison-Burch .....	7472
<i>A Two-Step Approach for Implicit Event Argument Detection</i>	
Zhisong Zhang, Xiang Kong, Zhengzhong Liu, Xuezhe Ma and Eduard Hovy .....	7479
<i>Machine Reading of Historical Events</i>	
Or Honovich, Lucas Torroba Hennigen, Omri Abend and Shay B. Cohen .....	7486
<i>Revisiting Unsupervised Relation Extraction</i>	
Thy Thy Tran, Phong Le and Sophia Ananiadou .....	7498
<i>SciREX: A Challenge Dataset for Document-Level Information Extraction</i>	
Sarthak Jain, Madeleine van Zuylen, Hannaneh Hajishirzi and Iz Beltagy .....	7506
<i>Contrastive Self-Supervised Learning for Commonsense Reasoning</i>	
Tassilo Klein and Moin Nabi .....	7517

<i>Do Transformers Need Deep Long-Range Memory?</i>	
Jack Rae and Ali Razavi .....	7524
<i>Improving Disentangled Text Representation Learning with Information-Theoretic Guidance</i>	
Pengyu Cheng, Martin Renqiang Min, Dinghan Shen, Christopher Malon, Yizhe Zhang, Yitong Li and Lawrence Carin .....	7530
<i>Understanding Advertisements with BERT</i>	
Kanika Kalra, Bhargav Kurma, Silpa Vadakkeveetil Sreelatha, Manasi Patwardhan and Shirish Karande .....	7542
<i>Non-Linear Instance-Based Cross-Lingual Mapping for Non-Isomorphic Embedding Spaces</i>	
Goran Glavaš and Ivan Vulić .....	7548
<i>Good-Enough Compositional Data Augmentation</i>	
Jacob Andreas .....	7556
<i>RAT-SQL: Relation-Aware Schema Encoding and Linking for Text-to-SQL Parsers</i>	
Bailin Wang, Richard Shin, Xiaodong Liu, Oleksandr Polozov and Matthew Richardson .....	7567
<i>Temporal Common Sense Acquisition with Minimal Supervision</i>	
Ben Zhou, Qiang Ning, Daniel Khashabi and Dan Roth .....	7579
<i>The Sensitivity of Language Models and Humans to Winograd Schema Perturbations</i>	
Mostafa Abdou, Vinit Ravishankar, Maria Barrett, Yonatan Belinkov, Desmond Elliott and Anders Søgaard .....	7590
<i>Temporally-Informed Analysis of Named Entity Recognition</i>	
Shruti Rijhwani and Daniel Preotiuc-Pietro .....	7605
<i>Towards Open Domain Event Trigger Identification using Adversarial Domain Adaptation</i>	
Aakanksha Naik and Carolyn Rose .....	7618
<i>CompGuessWhat?!: A Multi-task Evaluation Framework for Grounded Language Learning</i>	
Alessandro Suglia, Ioannis Konstas, Andrea Vanzo, Emanuele Bastianelli, Desmond Elliott, Stella Frank and Oliver Lemon .....	7625
<i>Cross-Modality Relevance for Reasoning on Language and Vision</i>	
Chen Zheng, Quan Guo and Parisa Kordjamshidi .....	7642
<i>Learning Web-based Procedures by Reasoning over Explanations and Demonstrations in Context</i>	
Shashank Srivastava, Oleksandr Polozov, Nebojsa Jojic and Christopher Meek .....	7652
<i>Multi-agent Communication meets Natural Language: Synergies between Functional and Structural Language Learning</i>	
Angeliki Lazaridou, Anna Potapenko and Olivier Tieleman .....	7663
<i>HAT: Hardware-Aware Transformers for Efficient Natural Language Processing</i>	
Hanrui Wang, Zhanghao Wu, Zhijian Liu, Han Cai, Ligeng Zhu, Chuang Gan and Song Han ..	7675
<i>Hard-Coded Gaussian Attention for Neural Machine Translation</i>	
Weiyou You, Simeng Sun and Mohit Iyyer .....	7689
<i>In Neural Machine Translation, What Does Transfer Learning Transfer?</i>	
Alham Fikri Aji, Nikolay Bogoychev, Kenneth Heafield and Rico Sennrich .....	7701

<i>Learning a Multi-Domain Curriculum for Neural Machine Translation</i>	
Wei Wang, Ye Tian, Jiquan Ngiam, Yinfei Yang, Isaac Caswell and Zarana Parekh . . . . .	7711
<i>Reducing Gender Bias in Neural Machine Translation as a Domain Adaptation Problem</i>	
Danielle Saunders and Bill Byrne . . . . .	7724
<i>Translationese as a Language in "Multilingual" NMT</i>	
Parker Riley, Isaac Caswell, Markus Freitag and David Grangier . . . . .	7737
<i>Unsupervised Domain Clusters in Pretrained Language Models</i>	
Roev Aharoni and Yoav Goldberg . . . . .	7747
<i>Using Context in Neural Machine Translation Training Objectives</i>	
Danielle Saunders, Felix Stahlberg and Bill Byrne . . . . .	7764
<i>Variational Neural Machine Translation with Normalizing Flows</i>	
Hendra Setiawan, Matthias Sperber, Udhyakumar Nallasamy and Matthias Paulik . . . . .	7771
<i>The Paradigm Discovery Problem</i>	
Alexander Erdmann, Micha Elsner, Shijie Wu, Ryan Cotterell and Nizar Habash . . . . .	7778
<i>Supervised Grapheme-to-Phoneme Conversion of Orthographic Schwas in Hindi and Punjabi</i>	
Aryaman Arora, Luke Gessler and Nathan Schneider . . . . .	7791
<i>Automated Evaluation of Writing – 50 Years and Counting</i>	
Beata Beigman Klebanov and Nitin Madnani . . . . .	7796
<i>Negated and Misprimed Probes for Pretrained Language Models: Birds Can Talk, But Cannot Fly</i>	
Nora Kassner and Hinrich Schütze . . . . .	7811
<i>On Forgetting to Cite Older Papers: An Analysis of the ACL Anthology</i>	
Marcel Bollmann and Desmond Elliott . . . . .	7819
<i>Returning the N to NLP: Towards Contextually Personalized Classification Models</i>	
Lucie Flek . . . . .	7828
<i>To Test Machine Comprehension, Start by Defining Comprehension</i>	
Jesse Dunietz, Greg Burnham, Akash Bharadwaj, Owen Rambow, Jennifer Chu-Carroll and Dave Ferrucci . . . . .	7839
<i>Gender Gap in Natural Language Processing Research: Disparities in Authorship and Citations</i>	
Saif M. Mohammad . . . . .	7860
<i>BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension</i>	
Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov and Luke Zettlemoyer . . . . .	7871
<i>BLEURT: Learning Robust Metrics for Text Generation</i>	
Thibault Sellam, Dipanjan Das and Ankur Parikh . . . . .	7881
<i>Distilling Knowledge Learned in BERT for Text Generation</i>	
Yen-Chun Chen, Zhe Gan, Yu Cheng, Jingzhou Liu and Jingjing Liu . . . . .	7893

<i>ESPRIT: Explaining Solutions to Physical Reasoning Tasks</i>	
Nazneen Fatema Rajani, Rui Zhang, Yi Chern Tan, Stephan Zheng, Jeremy Weiss, Aadit Vyas, Abhijit Gupta, Caiming Xiong, Richard Socher and Dragomir Radev .....	7906
<i>Iterative Edit-Based Unsupervised Sentence Simplification</i>	
Dhruv Kumar, Lili Mou, Lukasz Golab and Olga Vechtomova .....	7918
<i>Logical Natural Language Generation from Open-Domain Tables</i>	
Wenhu Chen, Jianshu Chen, Yu Su, Zhiyu Chen and William Yang Wang .....	7929
<i>Neural CRF Model for Sentence Alignment in Text Simplification</i>	
Chao Jiang, Mounica Maddela, Wuwei Lan, Yang Zhong and Wei Xu .....	7943
<i>One Size Does Not Fit All: Generating and Evaluating Variable Number of Keyphrases</i>	
Xingdi Yuan, Tong Wang, Rui Meng, Khushboo Thaker, Peter Brusilovsky, Daqing He and Adam Trischler .....	7961
<i>R<sup>3</sup>: Reverse, Retrieve, and Rank for Sarcasm Generation with Commonsense Knowledge</i>	
Tuhin Chakrabarty, Debanjan Ghosh, Smaranda Muresan and Nanyun Peng .....	7976
<i>Structural Information Preserving for Graph-to-Text Generation</i>	
Linfeng Song, Ante Wang, Jinsong Su, Yue Zhang, Kun Xu, Yubin Ge and Dong Yu .....	7987
<i>A Joint Neural Model for Information Extraction with Global Features</i>	
Ying Lin, Heng Ji, Fei Huang and Lingfei Wu .....	7999
<i>Document-Level Event Role Filler Extraction using Multi-Granularity Contextualized Encoding</i>	
Xinya Du and Claire Cardie .....	8010
<i>Exploiting the Syntax-Model Consistency for Neural Relation Extraction</i>	
Amir Pouran Ben Veyseh, Franck Dernoncourt, Dejing Dou and Thien Huu Nguyen .....	8021
<i>From English to Code-Switching: Transfer Learning with Strong Morphological Clues</i>	
Gustavo Aguilar and Thamar Solorio .....	8033
<i>Learning Interpretable Relationships between Entities, Relations and Concepts via Bayesian Structure Learning on Open Domain Facts</i>	
Jingyuan Zhang, Mingming Sun, Yue Feng and Ping Li .....	8045
<i>Multi-Sentence Argument Linking</i>	
Seth Ebner, Patrick Xia, Ryan Culkin, Kyle Rawlins and Benjamin Van Durme .....	8057
<i>Rationalizing Medical Relation Prediction from Corpus-level Statistics</i>	
Zhen Wang, Jennifer Lee, Simon Lin and Huan Sun .....	8078
<i>Sources of Transfer in Multilingual Named Entity Recognition</i>	
David Mueller, Nicholas Andrews and Mark Dredze .....	8093
<i>ZeroShotCeres: Zero-Shot Relation Extraction from Semi-Structured Webpages</i>	
Colin Lockard, Prashant Shiralkar, Xin Luna Dong and Hannaneh Hajishirzi .....	8105
<i>Soft Gazetteers for Low-Resource Named Entity Recognition</i>	
Shruti Rijhwani, Shuyan Zhou, Graham Neubig and Jaime Carbonell .....	8118
<i>A Prioritization Model for Suicidality Risk Assessment</i>	
Han-Chin Shing, Philip Resnik and Douglas Oard .....	8124

<i>CluHTM - Semantic Hierarchical Topic Modeling based on CluWords</i>	
Felipe Viegas, Washington Cunha, Christian Gomes, Antônio Pereira, Leonardo Rocha and Marcos Goncalves .....	8138
<i>Empower Entity Set Expansion via Language Model Probing</i>	
Yunyi Zhang, Jiaming Shen, Jingbo Shang and Jiawei Han .....	8151
<i>Feature Projection for Improved Text Classification</i>	
Qi Qin, Wenpeng Hu and Bing Liu .....	8161
<i>A negative case analysis of visual grounding methods for VQA</i>	
Robik Shrestha, Kushal Kafle and Christopher Kanan .....	8172
<i>History for Visual Dialog: Do we really need it?</i>	
Shubham Agarwal, Trung Bui, Joon-Young Lee, Ioannis Konstas and Verena Rieser .....	8182
<i>Mapping Natural Language Instructions to Mobile UI Action Sequences</i>	
Yang Li, Jiacong He, Xin Zhou, Yuan Zhang and Jason Baldridge .....	8198
<i>TVQA+: Spatio-Temporal Grounding for Video Question Answering</i>	
Jie Lei, Licheng Yu, Tamara Berg and Mohit Bansal .....	8211
<i>Unsupervised Multimodal Neural Machine Translation with Pseudo Visual Pivoting</i>	
Po-Yao Huang, Junjie Hu, Xiaojun Chang and Alexander Hauptmann .....	8226
<i>A Multitask Learning Approach for Diacritic Restoration</i>	
Sawsan Alqahtani, Ajay Mishra and Mona Diab .....	8238
<i>Frugal Paradigm Completion</i>	
Alexander Erdmann, Tom Kenter, Markus Becker and Christian Schallhart .....	8248
<i>Improving Chinese Word Segmentation with Wordhood Memory Networks</i>	
Yuanhe Tian, Yan Song, Fei Xia, Tong Zhang and Yonggang Wang .....	8274
<i>Joint Chinese Word Segmentation and Part-of-speech Tagging via Two-way Attentions of Auto-analyzed Knowledge</i>	
Yuanhe Tian, Yan Song, Xiang Ao, Fei Xia, Xiaojun Quan, Tong Zhang and Yonggang Wang .....	8286
<i>Joint Diacritization, Lemmatization, Normalization, and Fine-Grained Morphological Tagging</i>	
Nasser Zalmout and Nizar Habash .....	8297
<i>Phonetic and Visual Priors for Decipherment of Informal Romanization</i>	
Maria Ryskina, Matthew R. Gormley and Taylor Berg-Kirkpatrick .....	8308
<i>Active Learning for Coreference Resolution using Discrete Annotation</i>	
Belinda Z. Li, Gabriel Stanovsky and Luke Zettlemoyer .....	8320
<i>Beyond Possession Existence: Duration and Co-Possession</i>	
Dhivya Chinnappa, Srikala Murugan and Eduardo Blanco .....	8332
<i>Don't Stop Pretraining: Adapt Language Models to Domains and Tasks</i>	
Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey and Noah A. Smith .....	8342
<i>Estimating Mutual Information Between Dense Word Embeddings</i>	
Vitalii Zhelezniak, Aleksandar Savkov and Nils Hammerla .....	8361

<i>Exploring Unexplored Generalization Challenges for Cross-Database Semantic Parsing</i>	
Alane Suhr, Ming-Wei Chang, Peter Shaw and Kenton Lee .....	8372
<i>Predicting the Focus of Negation: Model and Error Analysis</i>	
Md Mosharaf Hossain, Kathleen Hamilton, Alexis Palmer and Eduardo Blanco .....	8389
<i>Structured Tuning for Semantic Role Labeling</i>	
Tao Li, Parth Anand Jawale, Martha Palmer and Vivek Srikumar .....	8402
<i>TaBERT: Pretraining for Joint Understanding of Textual and Tabular Data</i>	
Pengcheng Yin, Graham Neubig, Wen-tau Yih and Sebastian Riedel .....	8413
<i>Universal Compositional Semantic Parsing</i>	
Elias Stengel-Eskin, Aaron Steven White, Sheng Zhang and Benjamin Van Durme .....	8427
<i>Unsupervised Cross-lingual Representation Learning at Scale</i>	
Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer and Veselin Stoyanov .....	8440
<i>A Generate-and-Rank Framework with Semantic Type Regularization for Biomedical Concept Normalization</i>	
Dongfang Xu, Zeyu Zhang and Steven Bethard .....	8452
<i>Hierarchical Entity Typing via Multi-level Learning to Rank</i>	
Tongfei Chen, Yunmo Chen and Benjamin Van Durme .....	8465
<i>Multi-Domain Named Entity Recognition with Genre-Aware and Agnostic Inference</i>	
Jing Wang, Mayank Kulkarni and Daniel Preotiuc-Pietro .....	8476
<i>TXtract: Taxonomy-Aware Knowledge Extraction for Thousands of Product Categories</i>	
Giannis Karamanolakis, Jun Ma and Xin Luna Dong .....	8489
<i>TriggerNER: Learning with Entity Triggers as Explanations for Named Entity Recognition</i>	
Bill Yuchen Lin, Dong-Ho Lee, Ming Shen, Ryan Moreno, Xiao Huang, Prashant Shiralkar and Xiang Ren .....	8503
<i>Addressing Posterior Collapse with Mutual Information for Improved Variational Neural Machine Translation</i>	
Arya D. McCarthy, Xian Li, Jiatao Gu and Ning Dong .....	8512
<i>Balancing Training for Multilingual Neural Machine Translation</i>	
Xinyi Wang, Yulia Tsvetkov and Graham Neubig .....	8526
<i>Evaluating Robustness to Input Perturbations for Neural Machine Translation</i>	
Xing Niu, Prashant Mathur, Georgiana Dinu and Yaser Al-Onaizan .....	8538
<i>Parallel Corpus Filtering via Pre-trained Language Models</i>	
Boliang Zhang, Ajay Nagesh and Kevin Knight .....	8545
<i>Regularized Context Gates on Transformer for Machine Translation</i>	
Xintong Li, Lemao Liu, Rui Wang, Guoping Huang and Max Meng .....	8555
<i>A Multi-Perspective Architecture for Semantic Code Search</i>	
Rajarshi Haldar, Lingfei Wu, JinJun Xiong and Julia Hockenmaier .....	8563

<i>Automated Topical Component Extraction Using Neural Network Attention Scores from Source-based Essay Scoring</i>	
Haoran Zhang and Diane Litman .....	8569
<i>Clinical Concept Linking with Contextualized Neural Representations</i>	
Elliot Schumacher, Andriy Mulyar and Mark Dredze .....	8585
<i>DeSePtion: Dual Sequence Prediction and Adversarial Examples for Improved Fact-Checking</i>	
Christopher Hidey, Tuhin Chakrabarty, Tariq Alhindi, Siddharth Varia, Kriste Krstovski, Mona Diab and Smaranda Muresan .....	8593
<i>Let Me Choose: From Verbal Context to Font Selection</i>	
Amirreza Shirani, Franck Dernoncourt, Jose Echevarria, Paul Asente, Nedim Lipka and Tamar Solorio .....	8607
<i>Multi-Label and Multilingual News Framing Analysis</i>	
Afra Feyza Akyürek, Lei Guo, Randa Elanwar, Prakash Ishwar, Margrit Betke and Derry Tanti Wijaya .....	8614
<i>Predicting Performance for Natural Language Processing Tasks</i>	
Mengzhou Xia, Antonios Anastasopoulos, Ruochen Xu, Yiming Yang and Graham Neubig ..	8625
<i>ScriptWriter: Narrative-Guided Script Generation</i>	
Yutao Zhu, Ruihua Song, Zhicheng Dou, Jian-Yun NIE and Jin Zhou .....	8647
<i>Should All Cross-Lingual Embeddings Speak English?</i>	
Antonios Anastasopoulos and Graham Neubig .....	8658
<i>Smart To-Do: Automatic Generation of To-Do Items from Emails</i>	
Sudipto Mukherjee, Subhabrata Mukherjee, Marcello Hasegawa, Ahmed Hassan Awadallah and Ryen White .....	8680
<i>Are Natural Language Inference Models IMPPRESSive? Learning IMPLicature and PRESupposition</i>	
Paloma Jeretic, Alex Warstadt, Suvrat Bhooshan and Adina Williams .....	8690
<i>End-to-End Bias Mitigation by Modelling Biases in Corpora</i>	
Rabeeh Karimi Mahabadi, Yonatan Belinkov and James Henderson .....	8706
<i>Mind the Trade-off: Debiasing NLU Models without Degrading the In-distribution Performance</i>	
Prasetya Ajie Utama, Nafise Sadat Moosavi and Iryna Gurevych .....	8717
<i>NILE : Natural Language Inference with Faithful Natural Language Explanations</i>	
Sawan Kumar and Partha Talukdar .....	8730
<i>QuASE: Question-Answer Driven Sentence Encoding</i>	
Hangfeng He, Qiang Ning and Dan Roth .....	8743
<i>Towards Robustifying NLI Models Against Lexical Dataset Biases</i>	
Xiang Zhou and Mohit Bansal .....	8759
<i>Uncertain Natural Language Inference</i>	
Tongfei Chen, Zhengping Jiang, Adam Poliak, Keisuke Sakaguchi and Benjamin Van Durme	8772
<i>Extracting Headless MWEs from Dependency Parse Trees: Parsing, Tagging, and Joint Modeling Approaches</i>	
Tianze Shi and Lillian Lee .....	8780

<i>Revisiting Higher-Order Dependency Parsers</i>	
Erick Fonseca and André F. T. Martins .....	8795
<i>SeqVAT: Virtual Adversarial Training for Semi-Supervised Sequence Labeling</i>	
Luoxin Chen, Weitong Ruan, Xinyue Liu and Jianhua Lu .....	8801
<i>Treebank Embedding Vectors for Out-of-domain Dependency Parsing</i>	
Joachim Wagner, James Barry and Jennifer Foster .....	8812