

Findings

**Findings of the Association for
Computational Linguistics**

Findings of ACL: EMNLP 2020

16 – 20 November, 2020

©2020 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-952148-90-3

Preface by the Program Committee Co-Chairs

Welcome to the very first volume of *Findings of ACL*, a new publication designed to operate as a “companion” to conferences in the ACL stable. This first volume, *Findings of ACL: EMNLP 2020* comprises papers selected from those submitted to EMNLP 2020 which were not selected to appear at the main conference. Despite this, the Programme Committee recognised the value of these submissions, and deemed them to be strong work, worthy of publication.

Papers published in *Findings* count as full publications, however they are not assigned a presentation slot in the conference, but rather are published purely online in a separate volume in the ACL Anthology. There are a number of motivations for this new publication, from allowing timely work to be published quickly, to being more accepting of solid work, and helping to manage the increasing reviewing burden on the community.

This new publication option allowed us to accept many more worthy papers that would have otherwise been rejected from EMNLP. From the 3,359 submissions reviewed for EMNLP, a total of 752 were accepted to the main conference, and a further 520 papers were accepted to *Findings*. The papers accepted to *Findings* equates to 15.5% of the reviewed papers, giving an aggregate acceptance rate for EMNLP and *Findings* of 37.9%. We requested that authors withdraw their *Findings* papers if they preferred to resubmit their work to subsequent conferences. After withdrawals, 447 papers remained, of which 332 were long papers, and 115 short.

Papers were accepted to *Findings* based on their reviews and AC assessments, particularly whether reviewers found merit in the ideas and approaches, and found no fundamental issues in terms of the work’s motivation, theory, experimentation, analysis and ethical considerations. Where there was broad agreement of the robustness of the work, but the papers would otherwise have been rejected from EMNLP, instead these papers were accepted into *Findings*. Judgements as to the suitability of a paper to *Findings* was decided by the Senior Areas Chairs (SACs) and Programme Chairs (PCs), initially through SACs proposing a large suite of borderline papers and a cutoff for acceptance to the main conference. Final decisions were made for the main conference before considering papers for *Findings*. We are confident that this procedure did not disadvantage particular papers, or classes of papers. Instead it gave authors a publication option for papers that would otherwise have been rejected, and based on the fact that 86% of papers accepted to *Findings* were not withdrawn, there is clear community support for the endeavour.

We appreciate that *Findings* may have an effect on the EMNLP workshops through removing from circulation some papers that would have been submitted to these venues. For this reason we have provided a mechanism for authors of *Findings* papers to present their work at a workshop as a non-archival presentation. This was done as a match-making service between authors of *Findings* papers and workshops, resulting in 139 *Findings* papers being assigned workshop presentations.

The reviewing process for *Findings* is largely the same as for the main conference, and accordingly we wish to thank all involved in EMNLP 2020 for their efforts, as detailed in the *Preface to the Proceedings of EMNLP 2020*. We would like to specifically thank:

- Tim Baldwin, the architect of *Findings*, and all the members of the ACL Reviewing Committee who helped to refine the proposal;
- SACs for making paper recommendation decisions for *Findings*;
- Jing Li and Lemao Liu, the Publication Chairs responsible for *Findings* papers;
- The workshop chairs, Jackie Cheung and Lonneke van der Plas, and the coordinators of the many workshops, who helped in providing a venue for *Findings* authors to present their work;

- The many participants in the ACL 2020 Reviewing meeting, and others who provided their constructive feedback on this publication.

We hope that *Findings* will continue to serve as a companion to future conferences, and become an important venue for excellent, widely-read, and highly cited work in NLP.

EMNLP 2020 Program Co-Chairs

Trevor Cohn, University of Melbourne, Australia

Yulan He, University of Warwick, UK

Yang Liu, Amazon – Alexa AI, USA

Table of Contents

<i>Fully Quantized Transformer for Machine Translation</i> Gabriele Prato, Ella Charlaix and Mehdi Rezagholizadeh	1
<i>Summarizing Chinese Medical Answer with Graph Convolution Networks and Question-focused Dual Attention</i> Ningyu Zhang, Shumin Deng, Juan Li, xi chen, Wei Zhang and Huajun Chen	15
<i>Stay Hungry, Stay Focused: Generating Informative and Specific Questions in Information-Seeking Conversations</i> Peng Qi, Yuhao Zhang and Christopher D. Manning	25
<i>Adapting BERT for Word Sense Disambiguation with Gloss Selection Objective and Example Sentences</i> Boon Peng Yap, Andrew Koh and Eng Siong Chng	41
<i>Adversarial Text Generation via Sequence Contrast Discrimination</i> Ke Wang and Xiaojun Wan	47
<i>GRACE: Gradient Harmonized and Cascaded Labeling for Aspect-based Sentiment Analysis</i> Huaishao Luo, Lei Ji, Tianrui Li, Daxin Jiang and Nan Duan	54
<i>Reducing Sentiment Bias in Language Models via Counterfactual Evaluation</i> Po-Sen Huang, Huan Zhang, Ray Jiang, Robert Stanforth, Johannes Welbl, Jack Rae, Vishal Maini, Dani Yogatama and Pushmeet Kohli	65
<i>Improving Text Understanding via Deep Syntax-Semantics Communication</i> Hao Fei, Yafeng Ren and Donghong Ji	84
<i>GRUEN for Evaluating Linguistic Quality of Generated Text</i> Wanzheng Zhu and Suma Bhat	94
<i>A Greedy Bit-flip Training Algorithm for Binarized Knowledge Graph Embeddings</i> Katsuhiko Hayashi, Koki Kishimoto and Masashi Shimbo	109
<i>Difference-aware Knowledge Selection for Knowledge-grounded Conversation Generation</i> Chujie Zheng, Yunbo Cao, Daxin Jiang and Minlie Huang	115
<i>An Attentive Recurrent Model for Incremental Prediction of Sentence-final Verbs</i> Wenyan Li, Alvin Grissom II and Jordan Boyd-Graber	126
<i>Transformer-GCRF: Recovering Chinese Dropped Pronouns with General Conditional Random Fields</i> Jingxuan Yang, Kerui Xu, Jun Xu, Si Li, Sheng GAO, Jun Guo, Ji-Rong Wen and Nianwen Xue	137
<i>Neural Speed Reading Audited</i> Anders Søgaard	148
<i>Converting the Point of View of Messages Spoken to Virtual Assistants</i> Gunhee Lee, Vera Zu, Sai Srujana Buddi, Dennis Liang, Purva Kulkarni and Jack FitzGerald ..	154
<i>Robustness to Modification with Shared Words in Paraphrase Identification</i> Zhouxing Shi and Minlie Huang	164

<i>Few-shot Natural Language Generation for Task-Oriented Dialog</i>	
Baolin Peng, Chenguang Zhu, Chunyuan Li, Xiujun Li, Jinchao Li, Michael Zeng and Jianfeng Gao	172
<i>Mimic and Conquer: Heterogeneous Tree Structure Distillation for Syntactic NLP</i>	
Hao Fei, Yafeng Ren and Donghong Ji	183
<i>A Hierarchical Network for Abstractive Meeting Summarization with Cross-Domain Pretraining</i>	
Chenguang Zhu, Ruochen Xu, Michael Zeng and Xuedong Huang	194
<i>Active Testing: An Unbiased Evaluation Method for Distantly Supervised Relation Extraction</i>	
Pengshuai Li, Xinsong Zhang, Weijia Jia and Wei Zhao	204
<i>Semantic Matching for Sequence-to-Sequence Learning</i>	
Ruiyi Zhang, Changyou Chen, Xinyuan Zhang, Ke Bai and Lawrence Carin	212
<i>How Decoding Strategies Affect the Verifiability of Generated Text</i>	
Luca Massarelli, Fabio Petroni, Aleksandra Piktus, Myle Ott, Tim Rocktäschel, Vassilis Plachouras, Fabrizio Silvestri and Sebastian Riedel	223
<i>Minimize Exposure Bias of Seq2Seq Models in Joint Entity and Relation Extraction</i>	
Ranran Haoran Zhang, Qianying Liu, Aysa Xuemo Fan, Heng Ji, Daojian Zeng, Fei Cheng, Daisuke Kawahara and Sadao Kurohashi	236
<i>Gradient-based Analysis of NLP Models is Manipulable</i>	
Junlin Wang, Jens Tuyls, Eric Wallace and Sameer Singh	247
<i>Pretrain-KGE: Learning Knowledge Representation from Pretrained Language Models</i>	
Zhiyuan Zhang, Xiaoqian Liu, Yi Zhang, Qi Su, Xu SUN and Bin He	259
<i>A Self-Refinement Strategy for Noise Reduction in Grammatical Error Correction</i>	
Masato Mita, Shun Kiyono, Masahiro Kaneko, Jun Suzuki and Kentaro Inui	267
<i>Understanding tables with intermediate pre-training</i>	
Julian Eisenschlos, Syrine Krichene and Thomas Müller	281
<i>Enhance Robustness of Sequence Labelling with Masked Adversarial Training</i>	
Luoxin Chen, Xinyue Liu, Weitong Ruan and Jianhua Lu	297
<i>Multilingual Argument Mining: Datasets and Analysis</i>	
Orith Toledo-Ronen, Matan Orbach, Yonatan Bilu, Artem Spector and Noam Slonim	303
<i>Improving Grammatical Error Correction with Machine Translation Pairs</i>	
Wangchunshu Zhou, Tao Ge, Chang Mu, Ke Xu, Furu Wei and Ming Zhou	318
<i>Machines Getting with the Program: Understanding Intent Arguments of Non-Canonical Directives</i>	
Won Ik Cho, Youngki Moon, Sangwhan Moon, Seok Min Kim and Nam Soo Kim	329
<i>The RELX Dataset and Matching the Multilingual Blanks for Cross-Lingual Relation Classification</i>	
Abdullatif Köksal and Arzucan Özgür	340
<i>Control, Generate, Augment: A Scalable Framework for Multi-Attribute Text Generation</i>	
Giuseppe Russo, Nora Hollenstein, Claudiu Cristian Musat and Ce Zhang	351
<i>Open-Ended Visual Question Answering by Multi-Modal Domain Adaptation</i>	
Yiming Xu, Lin Chen, Zhongwei Cheng, Lixin Duan and Jiebo Luo	367

<i>Dual Low-Rank Multimodal Fusion</i>	
Tao Jin, Siyu Huang, Yingming Li and Zhongfei Zhang	377
<i>Contextual Modulation for Relation-Level Metaphor Identification</i>	
Omnia Zayed, John P. McCrae and Paul Buitelaar	388
<i>Context-aware Stand-alone Neural Spelling Correction</i>	
Xiangci Li, Hairong Liu and Liang Huang	407
<i>A Novel Workflow for Accurately and Efficiently Crowdsourcing Predicate Senses and Argument Labels</i>	
Youxuan Jiang, Huaiyu Zhu, Jonathan K. Kummerfeld, Yunyao Li and Walter Lasecki	415
<i>KorNLI and KorSTS: New Benchmark Datasets for Korean Natural Language Understanding</i>	
Jiyeon Ham, Yo Joong Choe, Kyubyong Park, Ilji Choi and Hyungjoon Soh	422
<i>Dialogue Generation on Infrequent Sentence Functions via Structured Meta-Learning</i>	
Yifan Gao, Piji Li, Wei Bi, Xiaojiang Liu, Michael Lyu and Irwin King	431
<i>Exploring Versatile Generative Language Model Via Parameter-Efficient Transfer Learning</i>	
Zhaojiang Lin, Andrea Madotto and Pascale Fung	441
<i>A Fully Hyperbolic Neural Model for Hierarchical Multi-Class Classification</i>	
Federico López and Michael Strube	460
<i>Claim Check-Worthiness Detection as Positive Unlabelled Learning</i>	
Dustin Wright and Isabelle Augenstein	476
<i>ConceptBert: Concept-Aware Representation for Visual Question Answering</i>	
François Gardères, Maryam Ziaeeafard, baptiste abeloos and Freddy Lecue	489
<i>Bootstrapping a Crosslingual Semantic Parser</i>	
Tom Sherborne, Yumo Xu and Mirella Lapata	499
<i>Revisiting Representation Degeneration Problem in Language Modeling</i>	
Zhong Zhang, Chongming Gao, Cong Xu, Rui Miao, Qinli Yang and Junming Shao	518
<i>The workweek is the best time to start a family – A Study of GPT-2 Based Claim Generation</i>	
Shai Gretz, Yonatan Bilu, Edo Cohen-Karlik and Noam Slonim	528
<i>Dynamic Data Selection for Curriculum Learning via Ability Estimation</i>	
John P. Lalor and hong yu	545
<i>Fixed Encoder Self-Attention Patterns in Transformer-Based Machine Translation</i>	
Alessandro Raganato, Yves Scherrer and Jörg Tiedemann	556
<i>ZEST: Zero-shot Learning from Text Descriptions using Textual Similarity and Visual Summarization</i>	
Tzuf Paz-Argaman, Reut Tsarfaty, Gal Chechik and Yuval Atzmon	569
<i>Few-Shot Multi-Hop Relation Reasoning over Knowledge Bases</i>	
Chuxu Zhang, Lu Yu, Mandana Saebi, Meng Jiang and Nitesh Chawla	580
<i>A structure-enhanced graph convolutional network for sentiment analysis</i>	
Fanyu Meng, Junlan Feng, danping yin, Si Chen and Min Hu	586
<i>PBoS: Probabilistic Bag-of-Subwords for Generalizing Word Embedding</i>	
Zhao Jinman, Shawn Zhong, Xiaomin Zhang and Yingyu Liang	596

<i>Interpretable Entity Representations through Large-Scale Typing</i> Yasumasa Onoe and Greg Durrett	612
<i>Empirical Studies of Institutional Federated Learning For Natural Language Processing</i> Xinghua Zhu, Jianzong Wang, Zhenhou Hong and Jing Xiao	625
<i>NeuReduce: Reducing Mixed Boolean-Arithmetic Expressions by Recurrent Neural Network</i> Weijie Feng, Binbin Liu, Dongpeng Xu, Qilong Zheng and Yun Xu	635
<i>From Language to Language-ish: How Brain-Like is an LSTM's Representation of Nonsensical Language Stimuli?</i> Maryam Hashemzadeh, Greta Kaufeld, Martha White, Andrea E. Martin and Alona Fyshe	645
<i>Revisiting Pre-Trained Models for Chinese Natural Language Processing</i> Yiming Cui, Wanxiang Che, Ting Liu, Bing Qin, Shijin Wang and Guoping Hu	657
<i>Cascaded Semantic and Positional Self-Attention Network for Document Classification</i> Juyong Jiang, Jie Zhang and Kai Zhang	669
<i>Toward Recognizing More Entity Types in NER: An Efficient Implementation using Only Entity Lexicons</i> Minlong Peng, Ruotian Ma, Qi Zhang, Lujun Zhao, Mengxi Wei, Changlong Sun and Xuanjing Huang	678
<i>From Disjoint Sets to Parallel Data to Train Seq2Seq Models for Sentiment Transfer</i> Paulo Cavalin, Marisa Vasconcelos, Marcelo Grave, Claudio Pinhanez and Victor Henrique Alves Ribeiro	689
<i>Learning to Stop: A Simple yet Effective Approach to Urban Vision-Language Navigation</i> Jiannan Xiang, Xin Wang and William Yang Wang	699
<i>Document Ranking with a Pretrained Sequence-to-Sequence Model</i> Rodrigo Nogueira, Zhiying Jiang, Ronak Pradeep and Jimmy Lin	708
<i>Pruning Redundant Mappings in Transformer Models via Spectral-Normalized Identity Prior</i> Zi Lin, Jeremiah Liu, Zi Yang, Nan Hua and Dan Roth	719
<i>Rethinking Self-Attention: Towards Interpretability in Neural Parsing</i> Khalil Mrini, Franck Dernoncourt, Quan Hung Tran, Trung Bui, Walter Chang and Ndapa Nakashole	731
<i>PolicyQA: A Reading Comprehension Dataset for Privacy Policies</i> Wasi Ahmad, Jianfeng Chi, Yuan Tian and Kai-Wei Chang	743
<i>A Linguistic Analysis of Visually Grounded Dialogues Based on Spatial Expressions</i> Takuma Udagawa, Takato Yamazaki and Akiko Aizawa	750
<i>Efficient Context and Schema Fusion Networks for Multi-Domain Dialogue State Tracking</i> Su Zhu, Jieyu Li, Lu Chen and Kai Yu	766
<i>Syntactic and Semantic-driven Learning for Open Information Extraction</i> Jialong Tang, Yaojie Lu, Hongyu Lin, Xianpei Han, Le Sun, Xinyan Xiao and Hua Wu	782
<i>Group-wise Contrastive Learning for Neural Dialogue Generation</i> Hengyi Cai, Hongshen Chen, Yonghao Song, Zhuoye Ding, Yongjun Bao, Weipeng Yan and Xiaofang Zhao	793

<i>E-BERT: Efficient-Yet-Effective Entity Embeddings for BERT</i> Nina Poerner, Ulli Waltinger and Hinrich Schütze	803
<i>A Multi-task Learning Framework for Opinion Triplet Extraction</i> Chen Zhang, Qiuchi Li, Dawei Song and benyou wang	819
<i>Event Extraction as Multi-turn Question Answering</i> Fayuan Li, Weihua Peng, Yuguang Chen, Quan Wang, Lu Pan, Yajuan Lyu and Yong Zhu	829
<i>Improving QA Generalization by Concurrent Modeling of Multiple Biases</i> Mingzhu Wu, Nafise Sadat Moosavi, Andreas Rücklé and Iryna Gurevych	839
<i>Actor-Double-Critic: Incorporating Model-Based Critic for Task-Oriented Dialogue Systems</i> Yen-chen Wu, Bo-Hsiang Tseng and Milica Gasic	854
<i>Controlled Hallucinations: Learning to Generate Faithfully from Noisy Data</i> Katja Filippova	864
<i>Sequential Span Classification with Neural Semi-Markov CRFs for Biomedical Abstracts</i> Kosuke Yamada, Tsutomu Hirao, Ryohei Sasano, Koichi Takeda and Masaaki Nagata	871
<i>Where to Submit? Helping Researchers to Choose the Right Venue</i> Konstantin Kobs, Tobias Koopmann, Albin Zehe, David Fernes, Philipp Krop and Andreas Hotho	878
<i>AirConcierge: Generating Task-Oriented Dialogue via Efficient Large-Scale Knowledge Retrieval</i> Chieh-Yang Chen, Pei-Hsin Wang, Shih-Chieh Chang, Da-Cheng Juan, Wei Wei and Jia-Yu Pan	884
<i>DocStruct: A Multimodal Method to Extract Hierarchy Structure in Document for General Form Understanding</i> Zilong Wang, Mingjie Zhan, xuebo liu and ding liang	898
<i>Pretrained Language Models for Dialogue Generation with Multiple Input Sources</i> Yu Cao, Wei Bi, Meng Fang and Dacheng Tao	909
<i>A Study in Improving BLEU Reference Coverage with Diverse Automatic Paraphrasing</i> Rachel Bawden, Biao Zhang, Lisa Yankovskaya, Andre Tättar and Matt Post	918
<i>Cross-lingual Alignment Methods for Multilingual BERT: A Comparative Study</i> Saurabh Kulshreshtha, Jose Luis Redondo Garcia and Ching-Yun Chang	933
<i>Hybrid Emoji-Based Masked Language Models for Zero-Shot Abusive Language Detection</i> Michele Corazza, Stefano Menini, Elena Cabrio, Sara Tonelli and Serena Villata	943
<i>SeNsER: Learning Cross-Building Sensor Metadata Tagger</i> Yang Jiao, Jiacheng Li, Jiaman Wu, Dezhi Hong, Rajesh Gupta and Jingbo Shang	950
<i>Persian Ezafe Recognition Using Transformers and Its Role in Part-Of-Speech Tagging</i> Ehsan Doostmohammadi, Mino Nassajian and Adel Rahimi	961
<i>Scene Graph Modification Based on Natural Language Commands</i> Xuanli He, Quan Hung Tran, Gholamreza Haffari, Walter Chang, Zhe Lin, Trung Bui, Franck Dernoncourt and Nhan Dam	972

<i>LiMiT: The Literal Motion in Text Dataset</i>	
Irene Manotas, Ngoc Phuoc An Vo and Vadim Sheinin	991
<i>Transition-based Parsing with Stack-Transformers</i>	
Ramón Fernandez Astudillo, Miguel Ballesteros, Tahira Naseem, Austin Blodgett and Radu Florian	
1001	
<i>Generative Data Augmentation for Commonsense Reasoning</i>	
Yiben Yang, Chaitanya Malaviya, Jared Fernandez, Swabha Swayamdipta, Ronan Le Bras, Ji-Ping Wang, Chandra Bhagavatula, Yejin Choi and Doug Downey	1008
<i>HybridQA: A Dataset of Multi-Hop Question Answering over Tabular and Textual Data</i>	
Wenhu Chen, Hanwen Zha, Zhiyu Chen, Wenhan Xiong, Hong Wang and William Yang Wang	1026
<i>PhoBERT: Pre-trained language models for Vietnamese</i>	
Dat Quoc Nguyen and Anh Tuan Nguyen	1037
<i>ESTeR: Combining Word Co-occurrences and Word Associations for Unsupervised Emotion Detection</i>	
Sujatha Das Gollapalli, Polina Rozenshtein and See-Kiong Ng	1043
<i>Make Templates Smarter: A Template Based Data2Text System Powered by Text Stitch Model</i>	
Bingfeng Luo, Zuo Bai, Kunfeng Lai and Jianping Shen	1057
<i>GCDST: A Graph-based and Copy-augmented Multi-domain Dialogue State Tracking</i>	
Peng Wu, Bowei Zou, Ridong Jiang and AiTi Aw	1063
<i>Incorporating Stylistic Lexical Preferences in Generative Language Models</i>	
Hrituraj Singh, Gaurav Verma and Balaji Vasan Srinivasan	1074
<i>Why do you think that? Exploring Faithful Sentence-Level Rationales Without Supervision</i>	
Max Glockner, Ivan Habernal and Iryna Gurevych	1080
<i>Semi-Supervised Learning for Video Captioning</i>	
Ke Lin, Zhuoxin Gan and Liwei WANG	1096
<i>Multi²OIE: Multilingual Open Information Extraction Based on Multi-Head Attention with BERT</i>	
Youngbin Ro, Yukyung Lee and Pilsung Kang	1107
<i>LGPSolver - Solving Logic Grid Puzzles Automatically</i>	
Elgun Jabrayilzade and Selma Tekir	1118
<i>Using the Past Knowledge to Improve Sentiment Classification</i>	
Qi Qin, Wenpeng Hu and Bing Liu	1124
<i>High-order Semantic Role Labeling</i>	
Zuchao Li, Hai Zhao, Rui Wang and Kevin Parnow	1134
<i>Undersensitivity in Neural Reading Comprehension</i>	
Johannes Welbl, Pasquale Minervini, Max Bartolo, Pontus Stenetorp and Sebastian Riedel ...	1152
<i>HyperText: Endowing FastText with Hyperbolic Geometry</i>	
Yudong Zhu, Di Zhou, Jinghui Xiao, Xin Jiang, Xiao Chen and Qun Liu	1166
<i>AutoETER: Automated Entity Type Representation for Knowledge Graph Embedding</i>	
Guanglin Niu, Bo Li, Yongfei Zhang, Shiliang Pu and Jingyang Li	1172

<i>Learning Robust and Multilingual Speech Representations</i>	
Kazuya Kawakami, Luyu Wang, Chris Dyer, Phil Blunsom and Aaron van den Oord	1182
<i>FQuAD: French Question Answering Dataset</i>	
Martin D’Hoffschmidt, Wacim Belblidia, Quentin Heinrich, Tom Brendlé and Maxime Vidal	1193
<i>Dynamic Semantic Matching and Aggregation Network for Few-shot Intent Detection</i>	
Hoang Nguyen, Chenwei Zhang, Congying Xia and Philip Yu	1209
<i>Quantifying the Contextualization of Word Representations with Semantic Class Probing</i>	
Mengjie Zhao, Philipp Dufter, Yadollah Yaghoobzadeh and Hinrich Schütze	1219
<i>Learning to Generate Clinically Coherent Chest X-Ray Reports</i>	
Justin Lovelace and Bobak Mortazavi	1235
<i>FELIX: Flexible Text Editing Through Tagging and Insertion</i>	
Jonathan Mallinson, Aliaksei Severyn, Eric Malmi and Guillermo Garrido	1244
<i>What Can We Do to Improve Peer Review in NLP?</i>	
Anna Rogers and Isabelle Augenstein	1256
<i>Unsupervised Relation Extraction from Language Models using Constrained Cloze Completion</i>	
Ankur Goswami, Akshata Bhat, Hadar Ohana and Theodoros Rekatsinas	1263
<i>Biomedical Event Extraction with Hierarchical Knowledge Graphs</i>	
Kung-Hsiang Huang, Mu Yang and Nanyun Peng	1277
<i>Language Generation via Combinatorial Constraint Satisfaction: A Tree Search Enhanced Monte-Carlo Approach</i>	
Maosen Zhang, Nan Jiang, Lei Li and Yexiang Xue	1286
<i>Examining the Ordering of Rhetorical Strategies in Persuasive Requests</i>	
Omar Shaikh, Jiaao Chen, Jon Saad-Falcon, Polo Chau and Diyi Yang	1299
<i>Evaluating Models’ Local Decision Boundaries via Contrast Sets</i>	
Matt Gardner, Yoav Artzi, Victoria Basmov, Jonathan Berant, Ben Bogin, Sihao Chen, Pradeep Dasigi, Dheeru Dua, Yanai Elazar, Ananth Gottumukkala, Nitish Gupta, Hannaneh Hajishirzi, Gabriel Ilharco, Daniel Khashabi, Kevin Lin, Jiangming Liu, Nelson F. Liu, Phoebe Mulcaire, Qiang Ning, Sameer Singh, Noah A. Smith, Sanjay Subramanian, Reut Tsarfaty, Eric Wallace, Ally Zhang and Ben Zhou	1307
<i>Parsing with Multilingual BERT, a Small Corpus, and a Small Treebank</i>	
Ethan C. Chau, Lucy H. Lin and Noah A. Smith	1324
<i>OptSLA: an Optimization-Based Approach for Sequential Label Aggregation</i>	
Nasim Sabetpour, Adithya Kulkarni and Qi Li	1335
<i>Optimizing Word Segmentation for Downstream Task</i>	
Tatsuya Hiraoka, Sho Takase, Kei Uchiumi, Atsushi Keyaki and Naoaki Okazaki	1341
<i>Dynamically Updating Event Representations for Temporal Relation Classification with Multi-category Learning</i>	
Fei Cheng, Masayuki Asahara, Ichiro Kobayashi and Sadao Kurohashi	1352
<i>A Compare Aggregate Transformer for Understanding Document-grounded Dialogue</i>	
Longxuan Ma, Wei-Nan Zhang, runxin sun and Ting Liu	1358

<i>TextHide: Tackling Data Privacy in Language Understanding Tasks</i> Yangsibo Huang, Zhao Song, Danqi Chen, Kai Li and Sanjeev Arora	1368
<i>Modeling Intra and Inter-modality Incongruity for Multi-Modal Sarcasm Detection</i> Hongliang Pan, Zheng Lin, Peng Fu, Yatao Qi and Weiping Wang	1383
<i>Investigating Transferability in Pretrained Language Models</i> Alex Tamkin, Trisha Singh, Davide Giovanardi and Noah Goodman	1393
<i>Improving Knowledge-Aware Dialogue Response Generation by Using Human-Written Prototype Dialogues</i> Sixing Wu, Ying Li, Dawei Zhang and Zhonghai Wu	1402
<i>Filtering before Iteratively Referring for Knowledge-Grounded Response Selection in Retrieval-Based Chatbots</i> Jia-Chen Gu, Zhenhua Ling, Quan Liu, Zhigang Chen and Xiaodan Zhu	1412
<i>Privacy-Preserving News Recommendation Model Learning</i> Tao Qi, Fangzhao Wu, Chuhan Wu, Yongfeng Huang and Xing Xie	1423
<i>exBERT: Extending Pre-trained Models with Domain-specific Vocabulary Under Constrained Training Resources</i> Wen Tai, H. T. Kung, Xin Dong, Marcus Comiter and Chang-Fu Kuo	1433
<i>Balancing via Generation for Multi-Class Text Classification Improvement</i> Naama Tepper, Esther Goldbraich, Naama Zwerdling, George Kour, Ateret Anaby Tavor and Boaz Carmeli	1440
<i>Conditional Neural Generation using Sub-Aspect Functions for Extractive News Summarization</i> Zhengyuan Liu, Ke Shi and Nancy Chen	1453
<i>Research Replication Prediction Using Weakly Supervised Learning</i> Tianyi Luo, xingyu li, Hainan Wang and Yang Liu	1464
<i>Open Domain Question Answering based on Text Enhanced Knowledge Graph with Hyperedge Infusion</i> Jiale Han, Bo Cheng and Xu Wang	1475
<i>Inexpensive Domain Adaptation of Pretrained Language Models: Case Studies on Biomedical NER and Covid-19 QA</i> Nina Poerner, Ulli Waltinger and Hinrich Schütze	1482
<i>Semantically Driven Sentence Fusion: Modeling and Evaluation</i> Eyal Ben-David, Orgad Keller, Eric Malmi, Idan Szpektor and Roi Reichart	1491
<i>Pseudo-Bidirectional Decoding for Local Sequence Transduction</i> Wangchunshu Zhou, Tao Ge and Ke Xu	1506
<i>Predicting Responses to Psychological Questionnaires from Participants' Social Media Posts and Question Text Embeddings</i> Huy Vu, Suhaib Abdurahman, Sudeep Bhatia and Lyle Ungar	1512
<i>Will it Unblend?</i> Yuval Pinter, Cassandra L. Jacobs and Jacob Eisenstein	1525

<i>CodeBERT: A Pre-Trained Model for Programming and Natural Languages</i>	
Zhangyin Feng, Daya Guo, Duyu Tang, Nan Duan, Xiaocheng Feng, Ming Gong, Linjun Shou, Bing Qin, Ting Liu, Daxin Jiang and Ming Zhou	1536
<i>StyleDGPT: Stylized Response Generation with Pre-trained Language Models</i>	
Ze Yang, wei wu, Can Xu, Xinnian Liang, jiaqi bai, Liran Wang, Wei Wang and Zhoujun Li .	1548
<i>Enhancing Automated Essay Scoring Performance via Fine-tuning Pre-trained Language Models with Combination of Regression and Ranking</i>	
Ruosong Yang, Jiannong Cao, Zhiyuan Wen, Youzheng Wu and Xiaodong He	1560
<i>Neural Dialogue State Tracking with Temporally Expressive Networks</i>	
Junfan Chen, Richong Zhang, Yongyi Mao and Jie Xu	1570
<i>Inferring about fraudulent collusion risk on Brazilian public works contracts in official texts using a Bi-LSTM approach</i>	
MARCOS LIMA, Roberta Silva, Felipe Lopes de Souza Mendes, Leonardo R. de Carvalho, Aleteia Araujo and Flavio de Barros Vidal	1580
<i>Data-to-Text Generation with Style Imitation</i>	
Shuai Lin, Wentao Wang, Zichao Yang, Xiaodan Liang, Frank F. Xu, Eric Xing and Zhiting Hu	1589
<i>Teaching Machine Comprehension with Compositional Explanations</i>	
Qinyuan Ye, Xiao Huang, Elizabeth Boschee and Xiang Ren	1599
<i>A Knowledge-Driven Approach to Classifying Object and Attribute Coreferences in Opinion Mining</i>	
Jiahua Chen, Shuai Wang, Sahisnu Mazumder and Bing Liu	1616
<i>SimAlign: High Quality Word Alignments Without Parallel Training Data Using Static and Contextualized Embeddings</i>	
Masoud Jalili Sabet, Philipp Dufter, François Yvon and Hinrich Schütze	1627
<i>TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification</i>	
Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke and Leonardo Neves	1644
<i>Octa: Omissions and Conflicts in Target-Aspect Sentiment Analysis</i>	
Zhe Zhang, Chung-Wei Hang and Munindar Singh	1651
<i>On the Language Neutrality of Pre-trained Multilingual Representations</i>	
Jindřich Libovický, Rudolf Rosa and Alexander Fraser	1663
<i>Cost-effective Selection of Pretraining Data: A Case Study of Pretraining BERT on Social Media</i>	
Xiang Dai, Sarvnaz Karimi, Ben Hachey and Cecile Paris	1675
<i>TopicBERT for Energy Efficient Document Classification</i>	
Yatin Chaudhary, Pankaj Gupta, Khushbu Saxena, Vivek Kulkarni, Thomas Runkler and Hinrich Schütze	1682
<i>Improving Constituency Parsing with Span Attention</i>	
Yuanhe Tian, Yan Song, Fei Xia and Tong Zhang	1691
<i>RecoBERT: A Catalog Language Model for Text-Based Recommendations</i>	
Itzik Malkiel, Oren Barkan, Avi Caciularu, Noam Razin, Ori Katz and Noam Koenigstein . . .	1704

<i>Multi-Agent Mutual Learning at Sentence-Level and Token-Level for Neural Machine Translation</i> Baohao Liao, Yingbo Gao and Hermann Ney	1715
<i>DomBERT: Domain-oriented Language Model for Aspect-based Sentiment Analysis</i> Hu Xu, Bing Liu, Lei Shu and Philip Yu	1725
<i>RMM: A Recursive Mental Model for Dialogue Navigation</i> Homero Roman Roman, Yonatan Bisk, Jesse Thomason, Asli Celikyilmaz and Jianfeng Gao .	1732
<i>Will This Idea Spread Beyond Academia? Understanding Knowledge Transfer of Scientific Concepts across Text Corpora</i> Hancheng Cao, Mengjie Cheng, Zhepeng Cen, Daniel McFarland and Xiang Ren	1746
<i>Recurrent Inference in Text Editing</i> Ning Shi, Ziheng Zeng, Haotian Zhang and Yichen Gong	1758
<i>An Empirical Exploration of Local Ordering Pre-training for Structured Prediction</i> Zhisong Zhang, Xiang Kong, Lori Levin and Eduard Hovy	1770
<i>Unsupervised Extractive Summarization by Pre-training Hierarchical Transformers</i> Shusheng Xu, Xingxing Zhang, Yi Wu, Furu Wei and Ming Zhou	1784
<i>Active Learning Approaches to Enhancing Neural Machine Translation</i> Yuekai Zhao, Haoran Zhang, Shuchang Zhou and Zhihua Zhang	1796
<i>AGIF: An Adaptive Graph-Interactive Framework for Joint Multiple Intent Detection and Slot Filling</i> Libo Qin, Xiao Xu, Wanxiang Che and Ting Liu	1807
<i>Continual Learning Long Short Term Memory</i> Xin Guo, Yu Tian, Qinghan Xue, Panos Lampropoulos, Steven Eliuk, Kenneth Barner and Xiaolong wang	1817
<i>CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning</i> Bill Yuchen Lin, Wangchunshu Zhou, Ming Shen, Pei Zhou, Chandra Bhagavatula, Yejin Choi and Xiang Ren	1823
<i>Constrained Decoding for Computationally Efficient Named Entity Recognition Taggers</i> Brian Lester, Daniel Pressel, Amy Hemmeter, Sagnik Ray Choudhury and Srinivas Bangalore	1841
<i>On the Potential of Lexico-logical Alignments for Semantic Parsing to SQL Queries</i> Tianze Shi, Chen Zhao, Jordan Boyd-Graber, Hal Daumé III and Lillian Lee	1849
<i>TED: A Pretrained Unsupervised Summarization Model with Theme Modeling and Denoising</i> Ziyi Yang, Chenguang Zhu, Robert Gmyr, Michael Zeng, Xuedong Huang and Eric Darve ...	1865
<i>Improving End-to-End Bangla Speech Recognition with Semi-supervised Training</i> Nafis Sadeq, Nafis Tahmid Chowdhury, Farhan Tanvir Utshaw, Shafayat Ahmed and Muhammad Abdullah Adnan	1875
<i>No Gestures Left Behind: Learning Relationships between Spoken Language and Freeform Gestures</i> Chaitanya Ahuja, Dong Won Lee, Ryo Ishii and Louis-Philippe Morency	1884
<i>UNIFIEDQA: Crossing Format Boundaries with a Single QA System</i> Daniel Khashabi, Sewon Min, Tushar Khot, Ashish Sabharwal, Oyvind Tafjord, Peter Clark and Hannaneh Hajishirzi	1896

<i>Robust and Interpretable Grounding of Spatial References with Relation Networks</i>	
Tsung-Yen Yang, Andrew Lan and Karthik Narasimhan	1908
<i>Pragmatic Issue-Sensitive Image Captioning</i>	
Allen Nie, Reuben Cohn-Gordon and Christopher Potts	1924
<i>PTUM: Pre-training User Model from Unlabeled User Behaviors via Self-supervision</i>	
Chuhan Wu, Fangzhao Wu, Tao Qi, Jianxun Lian, Yongfeng Huang and Xing Xie	1939
<i>Adversarial Subword Regularization for Robust Neural Machine Translation</i>	
Jungsoo Park, Mujeen Sung, Jinhyuk Lee and Jaewoo Kang	1945
<i>Learning Visual-Semantic Embeddings for Reporting Abnormal Findings on Chest X-rays</i>	
Jianmo Ni, Chun-Nan Hsu, Amilcare Gentili and Julian McAuley	1954
<i>SynET: Synonym Expansion using Transitivity</i>	
Jiale Yu, Yongliang Shen, Xinyin Ma, Chenghao Jia, Chen Chen and Weiming Lu	1961
<i>Scheduled DropHead: A Regularization Method for Transformer Models</i>	
Wangchunshu Zhou, Tao Ge, Furu Wei, Ming Zhou and Ke Xu	1971
<i>Multi-Turn Dialogue Generation in E-Commerce Platform with the Context of Historical Dialogue</i>	
WeiSheng Zhang, Kaisong Song, Yangyang Kang, Zhongqing Wang, Changlong Sun, Xiaozhong Liu, Shoushan Li, Min Zhang and Luo Si	1981
<i>Automatically Identifying Gender Issues in Machine Translation using Perturbations</i>	
Hila Gonen and Kellie Webster	1991
<i>Ruler: Data Programming by Demonstration for Document Labeling</i>	
Sara Evensen, Chang Ge and Cagatay Demiralp	1996
<i>Dual Reconstruction: a Unifying Objective for Semi-Supervised Neural Machine Translation</i>	
Weijia Xu, Xing Niu and Marine Carpuat	2006
<i>Focus-Constrained Attention Mechanism for CVAE-based Response Generation</i>	
Zhi Cui, Yanran Li, Jiayi ZHANG, Jianwei Cui, Chen Wei and Bin Wang	2021
<i>Chunk-based Chinese Spelling Check with Global Optimization</i>	
Zuyi Bao, Chen Li and Rui Wang	2031
<i>Multi-pretraining for Large-scale Text Classification</i>	
Kang-Min Kim, Bumsu Hyeon, Yeachan Kim, Jun-Hyung Park and SangKeun Lee	2041
<i>End-to-End Speech Recognition and Disfluency Removal</i>	
Paria Jamshid Lou and Mark Johnson	2051
<i>Characterizing the Value of Information in Medical Notes</i>	
Chao-Chun Hsu, Shantanu Karnwal, Sendhil Mullainathan, Ziad Obermeyer and Chenhao Tan	2062
<i>KLearn: Background Knowledge Inference from Summarization Data</i>	
Maxime Peyrard and Robert West	2073
<i>Extracting Chemical-Protein Interactions via Calibrated Deep Neural Network and Self-training</i>	
Dongha Choi and Hyunju Lee	2086

<i>Logic2Text: High-Fidelity Natural Language Generation from Logical Forms</i>	
Zhiyu Chen, Wenhu Chen, Hanwen Zha, Xiyu Zhou, Yunkai Zhang, Sairam Sundaresan and William Yang Wang	2096
<i>MedICaT: A Dataset of Medical Images, Captions, and Textual References</i>	
Sanjay Subramanian, Lucy Lu Wang, Ben Bogin, Sachin Mehta, Madeleine van Zuylen, Sravanthi Parasa, Sameer Singh, Matt Gardner and Hannaneh Hajishirzi	2112
<i>TSDG: Content-aware Neural Response Generation with Two-stage Decoding Process</i>	
Junsheng Kong, Zhicheng Zhong, Yi Cai, Xin Wu and Da Ren	2121
<i>Unsupervised Cross-Lingual Adaptation of Dependency Parsers Using CRF Autoencoders</i>	
Zhao Li and Kewei Tu	2127
<i>Diversify Question Generation with Continuous Content Selectors and Question Type Modeling</i>	
Zhen Wang, Siwei Rao, Jie Zhang, Zhen Qin, Guangjian Tian and Jun Wang	2134
<i>Participatory Research for Low-resourced Machine Translation: A Case Study in African Languages</i>	
Wilhelmina Nekoto, Vukosi Marivate, Tshinondiwa Matsila, Timi Fasubaa, Taiwo Fagbohunge, Solomon Oluwole Akinola, Shamsuddeen Muhammad, Salomon KABONGO KABENAMUALU, Salomey Osei, Freshia Sackey, Rubungo Andre Niyongabo, Ricky Macharm, Perez Ogayo, Orevaoghene Ahia, Musie Meressa Berhe, Mofetoluwa Adeyemi, Masabata Mokgesi-Seling, Lawrence Okegbemi, Laura Martinus, Kolawole Tajudeen, Kevin Degila, Kelechi Ogueji, Kathleen Siminyu, Julia Kreutzer, Jason Webster, Jamiil Toure Ali, Jade Abbott, Iroro Orife, Ignatius Ezeani, Idris Abdulkadir Dangana, Herman Kamper, Hady Elshar, Goodness Duru, ghollah kioko, Murhabazi Espoir, Elan van Biljon, Daniel Whitenack, Christopher Onyefuluchi, Chris Chinenye Emezue, Bonaventure F. P. Dossou, Blessing Sibanda, Blessing Bassey, Ayodele Olabiyi, Arshath Ramkilowan, Alp Öktem, Adewale Akinfaderin and Abdallah Bashir	2144
<i>ConveRT: Efficient and Accurate Conversational Representations from Transformers</i>	
Matthew Henderson, Iñigo Casanueva, Nikola Mrkšić, Pei-Hao Su, Tsung-Hsien Wen and Ivan Vulić	2161
<i>Computer Assisted Translation with Neural Quality Estimation and Automatic Post-Editing</i>	
Ke Wang, Jiayi Wang, Niyu Ge, Yangbin Shi, Yu Zhao and Kai Fan	2175
<i>Zero-Shot Rationalization by Multi-Task Transfer Learning from Question Answering</i>	
Po-Nien Kung, Tse-Hsuan Yang, Yi-Cheng Chen, Sheng-Siang Yin and Yun-Nung Chen	2187
<i>The Role of Reentrancies in Abstract Meaning Representation Parsing</i>	
Marco Damonte, Ida Szubert, Shay B. Cohen and Mark Steedman	2198
<i>Cross-Lingual Suicidal-Oriented Word Embedding toward Suicide Prevention</i>	
Daeun Lee, Soyoung Park, Jiwon Kang, Daejin Choi and Jinyoung Han	2208
<i>Service-oriented Text-to-SQL Parsing</i>	
Wangsu Hu and Jilei Tian	2218
<i>Reinforcement Learning with Imbalanced Dataset for Data-to-Text Medical Report Generation</i>	
Toru Nishino, Ryota Ozaki, Yohei Momoki, Tomoki Taniguchi, Ryuji Kano, Norihisa Nakano, Yuki Tagawa, Motoki Taniguchi, Tomoko Ohkuma and keigo nakamura	2223
<i>Reducing Quantity Hallucinations in Abstractive Summarization</i>	
Zheng Zhao, Shay B. Cohen and Bonnie Webber	2237

<i>Rethinking Topic Modelling: From Document-Space to Term-Space</i> Magnus Sahlgren	2250
<i>Sparse and Decorrelated Representations for Stable Zero-shot NMT</i> Bokyung Son and Sungwon Lyu	2260
<i>A Semi-supervised Approach to Generate the Code-Mixed Text using Pre-trained Encoder and Transfer Learning</i> Deepak Gupta, Asif Ekbal and Pushpak Bhattacharyya	2267
<i>BERT-MK: Integrating Graph Contextualized Knowledge into Pre-trained Language Models</i> Bin He, Di Zhou, Jinghui Xiao, Xin Jiang, Qun Liu, Nicholas Jing Yuan and Tong Xu	2281
<i>Recursive Top-Down Production for Sentence Generation with Latent Trees</i> Shawn Tan, Yikang Shen, Alessandro Sordoni, Aaron Courville and Timothy J. O'Donnell ..	2291
<i>Guided Dialogue Policy Learning without Adversarial Learning in the Loop</i> Ziming Li, Sungjin Lee, Baolin Peng, Jinchao Li, Julia Kiseleva, Maarten de Rijke, Shahin Shayan-deh and Jianfeng Gao	2308
<i>MultiDM-GCN: Aspect-guided Response Generation in Multi-domain Multi-modal Dialogue System using Graph Convolutional Network</i> Mauajama Firdaus, NIDHI THAKUR and Asif Ekbal	2318
<i>Edge-Enhanced Graph Convolution Networks for Event Detection with Syntactic Relation</i> Shiyao Cui, Bowen Yu, Tingwen Liu, Zhenyu Zhang, Xuebin Wang and Jinqiao Shi	2329
<i>Semi-supervised Formality Style Transfer using Language Model Discriminator and Mutual Information Maximization</i> Kunal Chawla and Diyi Yang	2340
<i>Differentially Private Representation for NLP: Formal Guarantee and An Empirical Study on Privacy and Fairness</i> Lingjuan Lyu, Xuanli He and Yitong Li	2355
<i>Helpful or Hierarchical? Predicting the Communicative Strategies of Chat Participants, and their Impact on Success</i> Farzana Rashid, Tommaso Fornaciari, Dirk Hovy, Eduardo Blanco and Fernando Vega-Redondo	2366
<i>Learning Knowledge Bases with Parameters for Task-Oriented Dialogue Systems</i> Andrea Madotto, Samuel Cahyawijaya, Genta Indra Winata, Yan Xu, Zihan Liu, Zhaojiang Lin and Pascale Fung	2372
<i>Adapting Open Domain Fact Extraction and Verification to COVID-FACT through In-Domain Language Modeling</i> Zhenghao Liu, Chenyan Xiong, Zhuyun Dai, Si Sun, Maosong Sun and Zhiyuan Liu	2395
<i>ProphetNet: Predicting Future N-gram for Sequence-to-Sequence Pre-training</i> Weizhen Qi, Yu Yan, Yeyun Gong, Dayiheng Liu, Nan Duan, Jiusheng Chen, Ruofei Zhang and Ming Zhou	2401
<i>DivGAN: Towards Diverse Paraphrase Generation via Diversified Generative Adversarial Network</i> Yue Cao and Xiaojun Wan	2411

<i>Plug-and-Play Conversational Models</i>	
Andrea Madotto, Etsuko Ishii, Zhaojiang Lin, Sumanth Dathathri and Pascale Fung	2422
<i>Event-Driven Learning of Systematic Behaviours in Stock Markets</i>	
Xianchao Wu	2434
<i>Learning Improvised Chatbots from Adversarial Modifications of Natural Language Feedback</i>	
Makesh Narsimhan Sreedhar, Kun Ni and Siva Reddy	2445
<i>Adapting Coreference Resolution to Twitter Conversations</i>	
Berfin Aktaş, Veronika Solopova, Annalena Kohnert and Manfred Stede	2454
<i>On Romanization for Model Transfer Between Scripts in Neural Machine Translation</i>	
Chantal Amrhein and Rico Sennrich	2461
<i>COSMIC: COMmonSense knowledge for eMotion Identification in Conversations</i>	
Deepanway Ghosal, Navonil Majumder, Alexander Gelbukh, Rada Mihalcea and Soujanya Poria	
2470	
<i>Improving Compositional Generalization in Semantic Parsing</i>	
Inbar Oren, Jonathan Herzig, Nitish Gupta, Matt Gardner and Jonathan Berant	2482
<i>Answer Span Correction in Machine Reading Comprehension</i>	
Revanth Gangi Reddy, Md Arafat Sultan, Efsun Sarioglu Kayi, Rong Zhang, Vittorio Castelli and	
Avi Sil	2496
<i>On the Interplay Between Fine-tuning and Sentence-level Probing for Linguistic Knowledge in Pre-trained Transformers</i>	
Marius Mosbach, Anna Khokhlova, Michael A. Hedderich and Dietrich Klakow	2502
<i>Zero-shot Entity Linking with Efficient Long Range Sequence Modeling</i>	
Zonghai Yao, Liangliang Cao and Huapu Pan	2517
<i>How Does Context Matter? On the Robustness of Event Detection with Context-Selective Mask Generalization</i>	
Jian Liu, Yubo Chen, Kang Liu, Yantao Jia and Zhicheng Sheng	2523
<i>Adaptive Feature Selection for End-to-End Speech Translation</i>	
Biao Zhang, Ivan Titov, Barry Haddow and Rico Sennrich	2533
<i>Abstractive Multi-Document Summarization via Joint Learning with Single-Document Summarization</i>	
Hanqi Jin and Xiaojun Wan	2545
<i>Blockwise Self-Attention for Long Document Understanding</i>	
Jiezhong Qiu, Hao Ma, Omer Levy, Wen-tau Yih, Sinong Wang and Jie Tang	2555
<i>Unsupervised Few-Bits Semantic Hashing with Implicit Topics Modeling</i>	
Fanghua Ye, Jarana Manotumrukta and Emine Yilmaz	2566
<i>Grid Tagging Scheme for Aspect-oriented Fine-grained Opinion Extraction</i>	
Zhen Wu, Chengcan Ying, Fei Zhao, Zhifang Fan, Xinyu Dai and Rui Xia	2576
<i>Learning Numeral Embedding</i>	
Chengyue Jiang, Zhonglin Nian, Kaihao Guo, Shanbo Chu, Yingong Zhao, Libin Shen and Kewei	
Tu	2586

<i>An Investigation of Potential Function Designs for Neural CRF</i> Zechuan Hu, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	2600
<i>Fast End-to-end Coreference Resolution for Korean</i> Cheoneum Park, Jamin Shin, Sungjoon Park, Joonho Lim and Changki Lee	2610
<i>Toward Stance-based Personas for Opinionated Dialogues</i> Thomas Scialom, Serra Sinem Tekirođlu, Jacopo Staiano and Marco Guerini	2625
<i>Hierarchical Pre-training for Sequence Labelling in Spoken Dialog</i> Emile Chapuis, Pierre Colombo, Matteo Manica, Matthieu Labeau and Chloé Clavel	2636
<i>Extending Multilingual BERT to Low-Resource Languages</i> Zihan Wang, Karthikeyan K, Stephen Mayhew and Dan Roth	2649
<i>Out-of-Sample Representation Learning for Knowledge Graphs</i> Marjan Albooyeh, Rishab Goel and Seyed Mehran Kazemi	2657
<i>Fine-Grained Grounding for Multimodal Speech Recognition</i> Tejas Srinivasan, Ramon Sanabria, Florian Metze and Desmond Elliott	2667
<i>Unsupervised Expressive Rules Provide Explainability and Assist Human Experts Grasping New Domains</i> Eyal Shnarch, Leshem Choshen, Guy Moshkovich, Ranit Aharonov and Noam Slonim	2678
<i>Textual Supervision for Visually Grounded Spoken Language Understanding</i> Bertrand Higy, Desmond Elliott and Grzegorz Chrupała	2698
<i>Universal Dependencies According to BERT: Both More Specific and More General</i> Tomasz Limisiewicz, David Mareček and Rudolf Rosa	2710
<i>Visual Objects As Context: Exploiting Visual Objects for Lexical Entailment</i> Masayasu Muraoka, Tetsuya Nasukawa and Bishwaranjan Bhattacharjee	2723
<i>Learning to Plan and Realize Separately for Open-Ended Dialogue Systems</i> Sashank Santhanam, Zhuo Cheng, Brodie Mather, Bonnie Dorr, Archana Bhatia, Bryanna Hebenstreit, Alan Zemel, Adam Dalton, Tomek Strzalkowski and Samira Shaikh	2736
<i>Be Different to Be Better! A Benchmark to Leverage the Complementarity of Language and Vision</i> Sandro Pezzelle, Claudio Greco, Greta Gandolfi, Eleonora Gualdoni and Raffaella Bernardi	2751
<i>Cross-Lingual Training of Neural Models for Document Ranking</i> Peng Shi, He Bai and Jimmy Lin	2768
<i>Improving Word Embedding Factorization for Compression Using Distilled Nonlinear Neural Decomposition</i> Vasileios Lioutas, Ahmad Rashid, Krtin Kumar, Md. Akmal Haidar and Mehdi Rezagholizadeh	2774
<i>PharmMT: A Neural Machine Translation Approach to Simplify Prescription Directions</i> Jiazhao Li, Corey Lester, Xinyan Zhao, Yuting Ding, Yun Jiang and V.G.Vinod Vydiswaran	2785
<i>LSTMs Compose—and Learn—Bottom-Up</i> Naomi Saphra and Adam Lopez	2797

<i>Natural Language Rationales with Full-Stack Visual Reasoning: From Pixels to Semantic Frames to Commonsense Graphs</i>	
Ana Marasović, Chandra Bhagavatula, Jae sung Park, Ronan Le Bras, Noah A. Smith and Yejin Choi	2810
<i>Corpora Evaluation and System Bias Detection in Multi-document Summarization</i>	
Alvin Dey, Tanya Chowdhury, Yash Kumar and Tanmoy Chakraborty	2830
<i>Graph-to-Tree Neural Networks for Learning Structured Input-Output Translation with Applications to Semantic Parsing and Math Word Problem</i>	
Shucheng Li, Lingfei Wu, Shiwei Feng, Fangli Xu, Fengyuan Xu and Sheng Zhong	2841
<i>Target Conditioning for One-to-Many Generation</i>	
marie-anne lachaux, Armand Joulin and Guillaume Lample	2853
<i>Can Pre-training help VQA with Lexical Variations?</i>	
Shailza Jolly and Shubham Kapoor	2863
<i>FENAS: Flexible and Expressive Neural Architecture Search</i>	
Ramakanth Pasunuru and Mohit Bansal	2869
<i>Inferring symmetry in natural language</i>	
Chelsea Tanchip, Lei Yu, Aotao Xu and Yang Xu	2877
<i>A Concise Model for Multi-Criteria Chinese Word Segmentation with Transformer Encoder</i>	
Xipeng Qiu, Hengzhi Pei, Hang Yan and Xuanjing Huang	2887
<i>LEGAL-BERT: The Muppets straight out of Law School</i>	
Ilias Chalkidis, Manos Fergadiotis, Prodromos Malakasiotis, Nikolaos Aletras and Ion Androutsopoulos	2898
<i>Enhancing Content Planning for Table-to-Text Generation with Data Understanding and Verification</i>	
Heng Gong, Wei Bi, Xiaocheng Feng, Bing Qin, Xiaojiang Liu and Ting Liu	2905
<i>Contextual Text Style Transfer</i>	
Yu Cheng, Zhe Gan, Yizhe Zhang, Oussama Elachqar, Dianqi Li and Jingjing Liu	2915
<i>DiPair: Fast and Accurate Distillation for Trillion-Scale Text Matching and Pair Modeling</i>	
Jiecao Chen, Liu Yang, Karthik Raman, Michael Bendersky, Jung-Jung Yeh, Yun Zhou, Marc Najork, Danyang Cai and Ehsan Emadzadeh	2925
<i>Cross-Lingual Dependency Parsing by POS-Guided Word Reordering</i>	
Lu Liu, Yi Zhou, Jianhan Xu, Xiaoqing Zheng, Kai-Wei Chang and Xuanjing Huang	2938
<i>Assessing Robustness of Text Classification through Maximal Safe Radius Computation</i>	
Emanuele La Malfa, Min Wu, Luca Laurenti, Benjie Wang, Anthony Hartshorn and Marta Kwiatkowska	2949
<i>Social Commonsense Reasoning with Multi-Head Knowledge Attention</i>	
Debjit Paul and Anette Frank	2969
<i>TurnGPT: a Transformer-based Language Model for Predicting Turn-taking in Spoken Dialog</i>	
Erik Ekstedt and Gabriel Skantze	2981
<i>A little goes a long way: Improving toxic language classification despite data scarcity</i>	
Mika Juuti, Tommi Gröndahl, Adrian Flanagan and N. Asokan	2991

<i>An Instance Level Approach for Shallow Semantic Parsing in Scientific Procedural Text</i> Daivik Swarup, Ahsaas Bajaj, Sheshera Mysore, Tim O’Gorman, Rajarshi Das and Andrew Mc- Callum	3010
<i>General Purpose Text Embeddings from Pre-trained Language Models for Scalable Inference</i> Jingfei Du, Myle Ott, Haoran Li, Xing Zhou and Veselin Stoyanov	3018
<i>Learning to Model and Ignore Dataset Bias with Mixed Capacity Ensembles</i> Christopher Clark, Mark Yatskar and Luke Zettlemoyer	3031
<i>Learning to Generalize for Sequential Decision Making</i> Xusen Yin, Ralph Weischedel and Jonathan May	3046
<i>Effective Crowd-Annotation of Participants, Interventions, and Outcomes in the Text of Clinical Trial Reports</i> Markus Zlabinger, Marta Sabou, Sebastian Hofstätter and Allan Hanbury	3064
<i>Adversarial Grammatical Error Correction</i> Vipul Raheja and Dimitris Alikaniotis	3075
<i>On Long-Tailed Phenomena in Neural Machine Translation</i> Vikas Raunak, Siddharth Dalmia, Vivek Gupta and Florian Metze	3088
<i>Knowing What You Know: Calibrating Dialogue Belief State Distributions via Ensembles</i> Carel van Niekerk, Michael Heck, Christian Geishauser, Hsien-chin Lin, Nurul Lubis, Marco Moresi and Milica Gasic	3096
<i>Domain Adversarial Fine-Tuning as an Effective Regularizer</i> Giorgos Vernikos, Katerina Margatina, Alexandra Chronopoulou and Ion Androutsopoulos ..	3103
<i>CLAR: A Cross-Lingual Argument Regularizer for Semantic Role Labeling</i> Ishan Jindal, Yunyao Li, Siddhartha Brahma and Huaiyu Zhu	3113
<i>Neutralizing Gender Bias in Word Embeddings with Latent Disentanglement and Counterfactual Gener- ation</i> Seungjae Shin, Kyungwoo Song, JoonHo Jang, Hyemi Kim, Weonyoung Joo and Il-Chul Moon	3126
<i>Towards Domain-Independent Text Structuring Trainable on Large Discourse Treebanks</i> Grigorii Guz and Giuseppe Carenini	3141
<i>Data Annealing for Informal Language Understanding Tasks</i> Jing Gu and Zhou Yu	3153
<i>A Multilingual View of Unsupervised Machine Translation</i> Xavier Garcia, Pierre Foret, Thibault Sellam and Ankur Parikh	3160
<i>An Evaluation Method for Diachronic Word Sense Induction</i> Ashjan Alsulaimani, Erwan Moreau and Carl Vogel	3171
<i>Integrating Task Specific Information into Pretrained Language Models for Low Resource Fine Tuning</i> Rui Wang, Shijing Si, Guoyin Wang, Lei Zhang, Lawrence Carin and Ricardo Henao	3181
<i>Efficient Transformer-based Large Scale Language Representations using Hardware-friendly Block Struc- tured Pruning</i> Bingbing Li, Zhenglun Kong, Tianyun Zhang, Ji Li, Zhengang Li, Hang Liu and Caiwen Ding	3187

<i>KoBE: Knowledge-Based Machine Translation Evaluation</i>	
Zorik Gekhman, Roei Aharoni, Genady Beryozkin, Markus Freitag and Wolfgang Macherey	3200
<i>Pushing the Limits of AMR Parsing with Self-Learning</i>	
Young-Suk Lee, Ramón Fernandez Astudillo, Tahira Naseem, Revanth Gangi Reddy, Radu Florian and Salim Roukos	3208
<i>Towards Zero-Shot Conditional Summarization with Adaptive Multi-Task Fine-Tuning</i>	
Travis Goodwin, Max Savery and Dina Demner-Fushman	3215
<i>Multilingual Knowledge Graph Completion via Ensemble Knowledge Transfer</i>	
Xuelu Chen, Muhao Chen, Changjun Fan, Ankith Uppunda, Yizhou Sun and Carlo Zaniolo	3227
<i>Towards Controllable Biases in Language Generation</i>	
Emily Sheng, Kai-Wei Chang, Prem Natarajan and Nanyun Peng	3239
<i>RobBERT: a Dutch RoBERTa-based Language Model</i>	
Pieter Delobelle, Thomas Winters and Bettina Berendt	3255
<i>Regularization of Distinct Strategies for Unsupervised Question Generation</i>	
Junmo Kang, Giwon Hong, Haritz Puerto San Roman and Sung-Hyon Myaeng	3266
<i>Graph-to-Graph Transformer for Transition-based Dependency Parsing</i>	
Alireza Mohammadshahi and James Henderson	3278
<i>WER we are and WER we think we are</i>	
Piotr Szymański, Piotr Żelasko, Mikolaj Morzy, Adrian Szymczak, Marzena Żyła-Hoppe, Joanna Banaszczyk, Lukasz Augustyniak, Jan Mizgajski and Yishay Carmiel	3290
<i>DeSMOG: Detecting Stance in Media On Global Warming</i>	
Yiwei Luo, Dallas Card and Dan Jurafsky	3296
<i>A Novel Challenge Set for Hebrew Morphological Disambiguation and Diacritics Restoration</i>	
Avi Shmidman, Joshua Guedalia, Shaltiel Shmidman, Moshe Koppel and Reut Tsarfaty	3316
<i>Improve Transformer Models with Better Relative Position Embeddings</i>	
zhiheng huang, Davis Liang, Peng Xu and Bing Xiang	3327
<i>A Sentiment-Controllable Topic-to-Essay Generator with Topic Knowledge Graph</i>	
Lin Qiao, Jianhao Yan, Fandong Meng, Zhendong Yang and Jie Zhou	3336
<i>What-if I ask you to explain: Explaining the effects of perturbations in procedural text</i>	
Dheeraj Rajagopal, Niket Tandon, Peter Clark, Bhavana Dalvi and Eduard Hovy	3345
<i>RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models</i>	
Samuel Gehman, Suchin Gururangan, Maarten Sap, Yejin Choi and Noah A. Smith	3356
<i>Improving Event Duration Prediction via Time-aware Pre-training</i>	
Zonglin Yang, Xinya Du, Alexander Rush and Claire Cardie	3370
<i>Composed Variational Natural Language Generation for Few-shot Intents</i>	
Congying Xia, Caiming Xiong, Philip Yu and Richard Socher	3379
<i>Literature Retrieval for Precision Medicine with Neural Matching and Faceted Summarization</i>	
Jiho Noh and Ramakanth Kavuluru	3389

<i>On the Importance of Adaptive Data Collection for Extremely Imbalanced Pairwise Tasks</i> Stephen Mussmann, Robin Jia and Percy Liang	3400
<i>A Dual-Attention Network for Joint Named Entity Recognition and Sentence Classification of Adverse Drug Events</i> Susmitha Wunnava, Xiao Qin, Tabassum Kakar, Xiangnan Kong and Elke Rundensteiner	3414
<i>BERT-kNN: Adding a kNN Search Component to Pretrained Language Models for Better QA</i> Nora Kassner and Hinrich Schütze	3424
<i>Identifying Spurious Correlations for Robust Text Classification</i> Zhao Wang and Aron Culotta	3431
<i>HoVer: A Dataset for Many-Hop Fact Extraction And Claim Verification</i> Yichen Jiang, Shikha Bordia, ZHENG ZHONG, Charles Dognin, Maneesh Singh and Mohit Bansal	3441
<i>Continual Learning for Natural Language Generation in Task-oriented Dialog Systems</i> Fei Mi, Liangwei Chen, Mengjie Zhao, Minlie Huang and Boi Faltings	3461
<i>UNQOVERing Stereotyping Biases via Underspecified Questions</i> Tao Li, Daniel Khashabi, Tushar Khot, Ashish Sabharwal and Vivek Srikumar	3475
<i>A Semantics-based Approach to Disclosure Classification in User-Generated Online Content</i> Chandan Akiti, Anna squicciarini and Sarah Rajtmajer	3490
<i>Mining Knowledge for Natural Language Inference from Wikipedia Categories</i> Mingda Chen, Zewei Chu, Karl Stratos and Kevin Gimpel	3500
<i>OCNLI: Original Chinese Natural Language Inference</i> Hai Hu, Kyle Richardson, Liang Xu, Lu Li, Sandra Kübler and Lawrence Moss	3512
<i>Margin-aware Unsupervised Domain Adaptation for Cross-lingual Text Labeling</i> Dejiao Zhang, Ramesh Nallapati, Henghui Zhu, Feng Nan, Cicero Nogueira dos Santos, Kathleen McKeown and Bing Xiang	3527
<i>Rethinking Supervised Learning and Reinforcement Learning in Task-Oriented Dialogue Systems</i> Ziming Li, Julia Kiseleva and Maarten de Rijke	3537
<i>What do we expect from Multiple-choice QA Systems?</i> Krunal Shah, Nitish Gupta and Dan Roth	3547
<i>Resource-Enhanced Neural Model for Event Argument Extraction</i> Jie Ma, Shuai Wang, Rishita Anubhai, Miguel Ballesteros and Yaser Al-Onaizan	3554
<i>Improving Target-side Lexical Transfer in Multilingual Neural Machine Translation</i> Luyu Gao, Xinyi Wang and Graham Neubig	3560
<i>Accurate polyglot semantic parsing with DAG grammars</i> Federico Fancellu, Ákos Kádár, Ran Zhang and Afsaneh Fazly	3567
<i>Approximation of Response Knowledge Retrieval in Knowledge-grounded Dialogue Generation</i> Wen Zheng, Natasa Milic-Frayling and Ke Zhou	3581
<i>Evaluating Factuality in Generation with Dependency-level Entailment</i> Tanya Goyal and Greg Durrett	3592

<i>Cross-Lingual Text Classification with Minimal Resources by Transferring a Sparse Teacher</i> Giannis Karamanolakis, Daniel Hsu and Luis Gravano	3604
<i>A Multi-Persona Chatbot for Hotline Counselor Training</i> Orianna Demasi, Yu Li and Zhou Yu	3623
<i>Narrative Text Generation with a Latent Discrete Plan</i> Harsh Jhamtani and Taylor Berg-Kirkpatrick	3637
<i>Graph Transformer Networks with Syntactic and Semantic Structures for Event Argument Extraction</i> Amir Pouran Ben Veyseh, Tuan Ngo Nguyen and Thien Huu Nguyen	3651
<i>The Box is in the Pen: Evaluating Commonsense Reasoning in Neural Machine Translation</i> Jie He, Tao Wang, Deyi Xiong and Qun Liu	3662
<i>Using Visual Feature Space as a Pivot Across Languages</i> Ziyan Yang, Leticia Pinto-Alva, Franck Dernoncourt and Vicente Ordonez	3673
<i>CDEvalSumm: An Empirical Study of Cross-Dataset Evaluation for Neural Summarization Systems</i> Yiran Chen, Pengfei Liu, Ming Zhong, Zi-Yi Dou, Danqing Wang, Xipeng Qiu and Xuanjing Huang	3679
<i>Attending to Long-Distance Document Context for Sequence Labeling</i> Matthew Jörke, Jon Gillick, Matthew Sims and David Bamman	3692
<i>Global Bootstrapping Neural Network for Entity Set Expansion</i> Lingyong Yan, Xianpei Han, Ben He and Le Sun	3705
<i>Document Classification for COVID-19 Literature</i> Bernal Jimenez Gutierrez, Jucheng Zeng, Dongdong Zhang, Ping Zhang and Yu Su	3715
<i>Adversarial Augmentation Policy Search for Domain and Cross-Lingual Generalization in Reading Comprehension</i> Adyasha Maharana and Mohit Bansal	3723
<i>Denoising Multi-Source Weak Supervision for Neural Text Classification</i> Wendi Ren, Yinghao Li, Hanting Su, David Kartchner, Cassie Mitchell and Chao Zhang	3739
<i>Dr. Summarize: Global Summarization of Medical Dialogue by Exploiting Local Structures.</i> Anirudh Joshi, Namit Katariya, Xavier Amatriain and Anitha Kannan	3755
<i>Generating Accurate Electronic Health Assessment from Medical Graph</i> Zhichao Yang and hong yu	3764
<i>Do Models of Mental Health Based on Social Media Data Generalize?</i> Keith Harrigan, Carlos Aguirre and Mark Dredze	3774
<i>Context Analysis for Pre-trained Masked Language Models</i> Yi-An Lai, Garima Lalwani and Yi Zhang	3789
<i>Controllable Text Generation with Focused Variation</i> Lei Shu, Alexandros Papangelis, Yi-Chia Wang, Gokhan Tur, Hu Xu, Zhaleh Feizollahi, Bing Liu and Piero Molino	3805

<i>Modeling Preconditions in Text with a Crowd-sourced Dataset</i>	
Heeyoung Kwon, Mahnaz Koupaee, Pratyush Singh, Gargi Sawhney, Anmol Shukla, Keerthi Kumar Kallur, Nathanael Chambers and Niranjana Balasubramanian	3818
<i>Reevaluating Adversarial Examples in Natural Language</i>	
John Morris, Eli Lifland, Jack Lanchantin, Yangfeng Ji and Yanjun Qi	3829
<i>Question Answering with Long Multiple-Span Answers</i>	
Ming Zhu, Aman Ahuja, Da-Cheng Juan, Wei Wei and Chandan K Reddy	3840
<i>Inserting Information Bottlenecks for Attribution in Transformers</i>	
Zhiying Jiang, Raphael Tang, Ji Xin and Jimmy Lin	3850
<i>Event-Related Bias Removal for Real-time Disaster Events</i>	
SALVADOR MEDINA MAZA, Evangelia Spiliopoulou, Eduard Hovy and Alexander Hauptmann	3858
<i>It's not a Non-Issue: Negation as a Source of Error in Machine Translation</i>	
Md Mosharaf Hossain, Antonios Anastasopoulos, Eduardo Blanco and Alexis Palmer	3869
<i>Incremental Text-to-Speech Synthesis with Prefix-to-Prefix Framework</i>	
Mingbo Ma, Baigong Zheng, Kaibo Liu, Renjie Zheng, Hairong Liu, Kainan Peng, Kenneth Church and Liang Huang	3886
<i>Joint Turn and Dialogue level User Satisfaction Estimation on Multi-Domain Conversations</i>	
Praveen Kumar Bodigutla, Aditya Tiwari, Spyros Matsoukas, Josep Valls-Vargas and Lazaros Polymenakos	3897
<i>ArraMon: A Joint Navigation-Assembly Instruction Interpretation Task in Dynamic Environments</i>	
Hyounghun Kim, Abhaysinh Zala, Graham Burri, Hao Tan and Mohit Bansal	3910
<i>Fluent and Low-latency Simultaneous Speech-to-Speech Translation with Self-adaptive Training</i>	
Renjie Zheng, Mingbo Ma, Baigong Zheng, Kaibo Liu, Jiahong Yuan, Kenneth Church and Liang Huang	3928
<i>Towards Context-Aware Code Comment Generation</i>	
Xiaohan Yu, Quzhe Huang, Zheng Wang, Yansong Feng and Dongyan Zhao	3938
<i>MCMH: Learning Multi-Chain Multi-Hop Rules for Knowledge Graph Reasoning</i>	
Lu Zhang, Mo Yu, Tian Gao and Yue Yu	3948
<i>Finding the Optimal Vocabulary Size for Neural Machine Translation</i>	
Thamme Gowda and Jonathan May	3955
<i>Weakly- and Semi-supervised Evidence Extraction</i>	
Danish Pruthi, Bhuwan Dhingra, Graham Neubig and Zachary C. Lipton	3965
<i>Making Information Seeking Easier: An Improved Pipeline for Conversational Search</i>	
Vaibhav Kumar and Jamie Callan	3971
<i>Generalizable and Explainable Dialogue Generation via Explicit Action Learning</i>	
Xinting Huang, Jianzhong Qi, Yu Sun and Rui Zhang	3981
<i>More Embeddings, Better Sequence Labelers?</i>	
Xinyu Wang, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	3992

<i>NLP Service APIs and Models for Efficient Registration of New Clients</i> Sahil Shah, Vihari Piratla, Soumen Chakrabarti and Sunita Sarawagi	4007
<i>Effects of Naturalistic Variation in Goal-Oriented Dialog</i> Jatin Ganhotra, Robert Moore, Sachindra Joshi and Kahini Wadhawan	4013
<i>Determining Event Outcomes: The Case of #fail</i> Srikala Murugan, Dhivya Chinnappa and Eduardo Blanco	4021
<i>WikiLingua: A New Benchmark Dataset for Cross-Lingual Abstractive Summarization</i> Faisal Ladhak, Esin Durmus, Claire Cardie and Kathleen McKeown	4034
<i>Adversarial Training for Code Retrieval with Question-Description Relevance Regularization</i> Jie Zhao and Huan Sun	4049
<i>Large Product Key Memory for Pretrained Language Models</i> Gyuwan Kim and Tae Hwan Jung	4060
<i>Temporal Reasoning in Natural Language Inference</i> Siddharth Vashishtha, Adam Poliak, Yash Kumar Lal, Benjamin Van Durme and Aaron Steven White	4070
<i>A Pilot Study of Text-to-SQL Semantic Parsing for Vietnamese</i> Anh Tuan Nguyen, Mai Hoang Dao and Dat Quoc Nguyen	4079
<i>STANDER: An Expert-Annotated Dataset for News Stance Detection and Evidence Retrieval</i> Costanza Conforti, Jakob Berndt, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd and Nigel Collier	4086
<i>An Empirical Methodology for Detecting and Prioritizing Needs during Crisis Events</i> M. Janina Sarol, Ly Dinh, Rezvaneh Rezapour, Chieh-Li Chin, Pingjing Yang and Jana Diesner	4102
<i>SupMMD: A Sentence Importance Model for Extractive Summarization using Maximum Mean Discrepancy</i> Umanga Bista, Alexander Mathews, Aditya Menon and Lexing Xie	4108
<i>Towards Low-Resource Semi-Supervised Dialogue Generation with Meta-Learning</i> Yi Huang, Junlan Feng, Shuo Ma, Xiaoyu Du and Xiaoting Wu	4123
<i>Connecting the Dots: A Knowledgeable Path Generator for Commonsense Question Answering</i> Peifeng Wang, Nanyun Peng, Filip Ilievski, Pedro Szekely and Xiang Ren	4129
<i>No Answer is Better Than Wrong Answer: A Reflection Model for Document Level Machine Reading Comprehension</i> Xuguang Wang, Linjun Shou, Ming Gong, Nan Duan and Daxin Jiang	4141
<i>Reference Language based Unsupervised Neural Machine Translation</i> Zuchao Li, Hai Zhao, Rui Wang, Masao Utiyama and Eiichiro Sumita	4151
<i>TinyBERT: Distilling BERT for Natural Language Understanding</i> Xiaoqi Jiao, Yichun Yin, Lifeng Shang, Xin Jiang, Xiao Chen, Linlin Li, Fang Wang and Qun Liu	4163
<i>Poison Attacks against Text Datasets with Conditional Adversarially Regularized Autoencoder</i> Alvin Chan, Yi Tay, Yew-Soon Ong and Aston Zhang	4175

<i>#TurkishTweets: A Benchmark Dataset for Turkish Text Correction</i>	
Asiye Tuba Koksak, Ozge Bozal, Emre Yürekli and Gizem Gezici	4190
<i>Assessing Human-Parity in Machine Translation on the Segment Level</i>	
Yvette Graham, Christian Federmann, Maria Eskevich and Barry Haddow	4199
<i>Multichannel Generative Language Model: Learning All Possible Factorizations Within and Across Channels</i>	
Harris Chan, Jamie Kiros and William Chan	4208
<i>Factorized Transformer for Multi-Domain Neural Machine Translation</i>	
Yongchao Deng, Hongfei Yu, Heng Yu, Xiangyu Duan and Weihua Luo	4221
<i>Improving Named Entity Recognition with Attentive Ensemble of Syntactic Information</i>	
Yuyang Nie, Yuanhe Tian, Yan Song, Xiang Ao and Xiang Wan	4231
<i>Query-Key Normalization for Transformers</i>	
Alex Henry, Prudhvi Raj Dachapally, Shubham Shantaram Pawar and Yuxuan Chen	4246
<i>Contract Discovery: Dataset and a Few-Shot Semantic Retrieval Challenge with Competitive Baselines</i>	
Łukasz Borchmann, Dawid Wisniewski, Andrzej Gretkowski, Izabela Kosmala, Dawid Jurkiewicz, Łukasz Szałkiewicz, Gabriela Pałka, Karol Kaczmarek, Agnieszka Kaliska and Filip Graliński . . .	4254
<i>Vocabulary Adaptation for Domain Adaptation in Neural Machine Translation</i>	
Shoetsu Sato, Jin Sakuma, Naoki Yoshinaga, Masashi Toyoda and Masaru Kitsuregawa	4269
<i>A Shared-Private Representation Model with Coarse-to-Fine Extraction for Target Sentiment Analysis</i>	
Peiqin Lin and Meng Yang	4280
<i>Detecting Media Bias in News Articles using Gaussian Bias Distributions</i>	
Wei-Fan Chen, Khalid Al Khatib, Benno Stein and Henning Wachsmuth	4290
<i>How Can Self-Attention Networks Recognize Dyck-n Languages?</i>	
Javid Ebrahimi, Dhruv Gelda and Wei Zhang	4301
<i>Training Flexible Depth Model by Multi-Task Learning for Neural Machine Translation</i>	
Qiang Wang, Tong Xiao and Jingbo Zhu	4307
<i>Looking inside Noun Compounds: Unsupervised Prepositional and Free Paraphrasing</i>	
Girishkumar Ponkiya, Rudra Murthy, Pushpak Bhattacharyya and Girish Palshikar	4313
<i>The birth of Romanian BERT</i>	
Stefan Dumitrescu, Andrei-Marius Avram and Sampo Pyysalo	4324
<i>BERT for Monolingual and Cross-Lingual Reverse Dictionary</i>	
Hang Yan, Xiaonan Li, Xipeng Qiu and bocaodeng	4329
<i>What's so special about BERT's layers? A closer look at the NLP pipeline in monolingual and multilingual models</i>	
Wietse de Vries, Andreas van Cranenburgh and Malvina Nissim	4339
<i>Leakage-Adjusted Simulatability: Can Models Generate Non-Trivial Explanations of Their Behavior in Natural Language?</i>	
Peter Hase, Shiyue Zhang, Harry Xie and Mohit Bansal	4351

<i>A Pointer Network Architecture for Joint Morphological Segmentation and Tagging</i> Amit Seker and Reut Tsarfaty	4368
<i>Beyond Language: Learning Commonsense from Images for Reasoning</i> Wanqing Cui, Yanyan Lan, Liang Pang, Jiafeng Guo and Xueqi Cheng	4379
<i>A BERT-based Distractor Generation Scheme with Multi-tasking and Negative Answer Training Strategies.</i> Ho-Lam Chung, Ying-Hong Chan and Yao-Chung Fan	4390
<i>How Effective is Task-Agnostic Data Augmentation for Pretrained Transformers?</i> Shayne Longpre, Yu Wang and Chris DuBois	4401
<i>Visually-Grounded Planning without Vision: Language Models Infer Detailed Plans from High-level Instructions</i> Peter Jansen	4412
<i>Consistent Response Generation with Controlled Specificity</i> Junya Takayama and Yuki Arase	4418
<i>Internal and external pressures on language emergence: least effort, object constancy and frequency</i> Diana Rodríguez Luna, Edoardo Maria Ponti, Dieuwke Hupkes and Elia Bruni	4428
<i>Parsing All: Syntax and Semantics, Dependencies and Spans</i> Junru Zhou, Zuchao Li and Hai Zhao	4438
<i>LIMIT-BERT : Linguistics Informed Multi-Task BERT</i> Junru Zhou, Zhuosheng Zhang, Hai Zhao and Shuailiang Zhang	4450
<i>Improving Limited Labeled Dialogue State Tracking with Self-Supervision</i> Chien-Sheng Wu, Steven C.H. Hoi and Caiming Xiong	4462
<i>On the Branching Bias of Syntax Extracted from Pre-trained Language Models</i> Huayang Li, Lema Liu, Guoping Huang and Shuming Shi	4473
<i>The Pragmatics behind Politics: Modelling Metaphor, Framing and Emotion in Political Discourse</i> Pere-Lluís Hugué Cabot, Verna Dankers, David Abadi, Agneta Fischer and Ekaterina Shutova	4479
<i>SMRT Chatbots: Improving Non-Task-Oriented Dialog with Simulated Multiple Reference Training</i> Huda Khayrallah and João Sedoc	4489
<i>PrivNet: Safeguarding Private Attributes in Transfer Learning for Recommendation</i> Guangneng Hu and Qiang Yang	4506
<i>Learning to Learn to Disambiguate: Meta-Learning for Few-Shot Word Sense Disambiguation</i> Nithin Holla, Pushkar Mishra, Helen Yannakoudakis and Ekaterina Shutova	4517
<i>An Empirical Investigation of Beam-Aware Training in Supertagging</i> Renato Negrinho, Matthew R. Gormley and Geoff Gordon	4534
<i>Improving Aspect-based Sentiment Analysis with Gated Graph Convolutional Networks and Syntax-based Regulation</i> Amir Pouran Ben Veyseh, Nasim Nouri, Franck Dernoncourt, Quan Hung Tran, Dejing Dou and Thien Huu Nguyen	4543

<i>Decoding Language Spatial Relations to 2D Spatial Arrangements</i> Gorjan Radevski, Guillem Collell, Marie-Francine Moens and Tinne Tuytelaars	4549
<i>The Dots Have Their Values: Exploiting the Node-Edge Connections in Graph-based Neural Models for Document-level Relation Extraction</i> Hieu Minh Tran, Minh Trung Nguyen and Thien Huu Nguyen	4561
<i>Why and when should you pool? Analyzing Pooling in Recurrent Architectures</i> Pratyush Maini, Keshav Kolluru, Danish Pruthi and	4568
<i>Structural and Functional Decomposition for Personality Image Captioning in a Communication Game</i> Minh Thu Nguyen, Duy Phung, Minh Hoai and Thien Huu Nguyen	4587
<i>Long Document Ranking with Query-Directed Sparse Transformer</i> Jyun-Yu Jiang, Chenyan Xiong, Chia-Jung Lee and Wei Wang	4594
<i>Visuo-Linguistic Question Answering (VLQA) Challenge</i> Shailaja Keyur Sampat, Yezhou Yang and Chitta Baral	4606
<i>Byte Pair Encoding is Suboptimal for Language Model Pretraining</i> Kaj Bostrom and Greg Durrett	4617
<i>Exploring BERT's Sensitivity to Lexical Cues using Tests from Semantic Priming</i> Kanishka Misra, Allyson Ettinger and Julia Rayz	4625
<i>Multi-hop Question Generation with Graph Convolutional Network</i> Dan Su, Yan Xu, Wenliang Dai, Ziwei Ji, Tiezheng Yu and Pascale Fung	4636
<i>MMFT-BERT: Multimodal Fusion Transformer with BERT Encodings for Visual Question Answering</i> Aisha Urooj, Amir Mazaheri, Niels Da vitoria lobo and Mubarak Shah	4648
<i>Thinking Like a Skeptic: Defeasible Inference in Natural Language</i> Rachel Rudinger, Vered Shwartz, Jena D. Hwang, Chandra Bhagavatula, Maxwell Forbes, Ronan Le Bras, Noah A. Smith and Yejin Choi	4661
<i>Guiding Attention for Self-Supervised Learning with Transformers</i> Ameet Deshpande and Karthik Narasimhan	4676
<i>Language-Conditioned Feature Pyramids for Visual Selection Tasks</i> Taichi Iki and Akiko Aizawa	4687
<i>Learning to Classify Events from Human Needs Category Descriptions</i> Haibo Ding and Zhe Feng	4698
<i>Automatic Term Name Generation for Gene Ontology: Task and Dataset</i> Yanjian Zhang, Qin Chen, Yiteng Zhang, Zhongyu Wei, Yixu Gao, Jiajie Peng, Zengfeng Huang, Weijian Sun and Xuanjing Huang	4705
<i>Compressing Transformer-Based Semantic Parsing Models using Compositional Code Embeddings</i> Prafull Prakash, Saurabh Kumar Shashidhar, Wenlong Zhao, Subendhu Rongali, Haidar Khan and Michael Kayser	4711
<i>BERT-QE: Contextualized Query Expansion for Document Re-ranking</i> Zhi Zheng, Kai Hui, Ben He, Xianpei Han, Le Sun and Andrew Yates	4718

<i>ZEN: Pre-training Chinese Text Encoder Enhanced by N-gram Representations</i> Shizhe Diao, Jiaxin Bai, Yan Song, Tong Zhang and Yonggang Wang	4729
<i>Finding Friends and Flipping Frenemies: Automatic Paraphrase Dataset Augmentation Using Graph Theory</i> Hannah Chen, Yangfeng Ji and David Evans	4741
<i>Probabilistic Case-based Reasoning for Open-World Knowledge Graph Completion</i> Rajarshi Das, Ameya Godbole, Nicholas Monath, Manzil Zaheer and Andrew McCallum	4752
<i>TLDR: Extreme Summarization of Scientific Documents</i> Isabel Cachola, Kyle Lo, Arman Cohan and Daniel Weld	4766
<i>Tri-Train: Automatic Pre-Fine Tuning between Pre-Training and Fine-Tuning for SciNER</i> Qingkai Zeng, Wenhao Yu, Mengxia Yu, Tianwen Jiang, Tim Weninger and Meng Jiang	4778
<i>Hierarchical Region Learning for Nested Named Entity Recognition</i> Xinwei Long, Shuzi Niu and Yucheng Li	4788
<i>Understanding User Resistance Strategies in Persuasive Conversations</i> Youzhi Tian, Weiyang Shi, Chen Li and Zhou Yu	4794
<i>On the Sub-layer Functionalities of Transformer Decoder</i> Yilin Yang, Longyue Wang, Shuming Shi, Prasad Tadepalli, Stefan Lee and Zhaopeng Tu	4799
<i>Extremely Low Bit Transformer Quantization for On-Device Neural Machine Translation</i> Insoo Chung, Byeongwook Kim, Yoonjung Choi, Se Jung Kwon, Yongkweon Jeon, Baeseong Park, Sangha Kim and Dongsoo Lee	4812
<i>Robust Backed-off Estimation of Out-of-Vocabulary Embeddings</i> Nobukazu Fukuda, Naoki Yoshinaga and Masaru Kitsuregawa	4827
<i>Exploiting Unsupervised Data for Emotion Recognition in Conversations</i> Wenxiang Jiao, Michael Lyu and Irwin King	4839
<i>Tensorized Embedding Layers</i> Oleksii Hrinchuk, Valentin Khrulkov, Leyla Mirvakhabova, Elena Orlova and Ivan Oseledets	4847
<i>Speaker or Listener? The Role of a Dialog Agent</i> Yafei Liu, Hongjin Qian, Hengpeng Xu and Jinmao Wei	4861
<i>Bridging Textual and Tabular Data for Cross-Domain Text-to-SQL Semantic Parsing</i> Xi Victoria Lin, Richard Socher and Caiming Xiong	4870
<i>Do Language Embeddings capture Scales?</i> Xikun Zhang, Deepak Ramachandran, Ian Tenney, Yanai Elazar and Dan Roth	4889
<i>Paraphrasing vs Coreferring: Two Sides of the Same Coin</i> Yehudit Meged, Avi Caciularu, Vered Shwartz and Ido Dagan	4897
<i>Active Sentence Learning by Adversarial Uncertainty Sampling in Discrete Space</i> Dongyu Ru, Jiangtao Feng, Lin Qiu, Hao Zhou, Mingxuan Wang, Weinan Zhang, Yong Yu and Lei Li	4908

<i>Coming to Terms: Automatic Formation of Neologisms in Hebrew</i> Moran Mizrahi, Stav Yardeni Seelig and Dafna Shahaf	4918
<i>Dual Inference for Improving Language Understanding and Generation</i> Shang-Yu Su, Yung-Sung Chuang and Yun-Nung Chen	4930
<i>Joint Intent Detection and Entity Linking on Spatial Domain Queries</i> lei zhang, Runze Wang, Jingbo Zhou, Jingsong Yu, Zhenhua Ling and Hui Xiong	4937
<i>IndicNLPsuite: Monolingual Corpora, Evaluation Benchmarks and Pre-trained Multilingual Language Models for Indian Languages</i> Divyanshu Kakwani, Anoop Kunchukuttan, Satish Golla, Gokul N.C., Avik Bhattacharyya, Mitesh M. Khapra and Pratyush Kumar	4948
<i>Weakly-Supervised Modeling of Contextualized Event Embedding for Discourse Relations</i> I-Ta Lee, Maria Leonor Pacheco and Dan Goldwasser	4962
<i>Enhancing Generalization in Natural Language Inference by Syntax</i> Qi He, Han Wang and Yue Zhang	4973

