Findings

Findings of the Association for Computational Linguistics

Findings of ACL: EMNLP 2021

November 7–11, 2021
Welcome to the *Findings* of ACL: EMNLP 2021! To continue the success of *Findings* of ACL series, we followed EMNLP 2020's initiative to produce this accompanying volume, for papers that narrowly missed acceptance to the main conference, but were judged to be solid, well-executed research, and worthy of publication. Out of the 3,600 submissions under review for EMNLP 2021, 445 papers were invited to be included in the *Findings*. 26 papers declined the offer, leading to 419 (300 long and 119 short) papers to be published in the *Findings* of ACL: EMNLP 2021. The papers accepted to *Findings* equates to 11.6% of the reviewed papers, giving an aggregate acceptance rate for EMNLP and *Findings* of 34.9%. In addition, we also accepted 5 long papers to the *Findings*, through the first ACL Rolling Review (ARR) pilot at EMNLP 2021. These papers are published in this volume as any other *Findings* paper.

Papers published in *Findings* of ACL count as full publications. They are not assigned a presentation slot in the main conference, but rather are published 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. To increase the visibility of the *Findings* papers, the authors are requested to submit a 6 (for long papers) or 3 minutes (for short papers) pre-recorded video presentation to be included in the virtual conference, hosted by Underline. Our workshop chairs also helped to pair *Findings* papers with EMNLP 2021 workshops, and as a result, around 100 *Findings* papers will be presented at those workshops.

The reviewing process for *Findings* is largely the same as for the main conference and accordingly we wish to thank all involved in EMNLP 2021 for their efforts, as detailed in the Preface to the Proceedings of EMNLP 2021. We would like to specifically thank:

- The whole **Program Committee** for reviewing the submissions, and in particular, the **Senior Area Chairs** for making recommendation on final paper decisions for *Findings*;
- The **Ethics Committee**, chaired by Margot Mieskes and Chris Potts, for their hard work to ensure that all the conditionally accepted *Findings* papers have addressed the ethical issues appropriately;
- **Past *ACL* PCs**, including Trevor Cohn, Yulan He and Yang Liu (EMNLP 2020), and Fei Xia, Wenjie Li, Roberto Navigli (ACL-IJCNLP 2021) for sharing their experience on *Findings* papers;
- **Findings Chairs** Gabriel Stanovsky and Tim Van de Cruys, for their tremendous effort in making the volume of *Findings of ACL*: EMNLP 2021;
- **Workshop Chairs** Parisa Kordjamshidi and Minlie Huang, for connecting *Findings* paper authors with workshop organizers for possible presentations.

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.

*Xuanjing Huang*, Fudan University, China  
*Lucia Specia*, Imperial College London, UK  
*Scott Wen-tao Yih*, Facebook, USA

EMNLP 2021 Program Committee Co-Chairs
# Table of Contents

**K-PLUG: Knowledge-injected Pre-trained Language Model for Natural Language Understanding and Generation in E-Commerce**  
Song Xu, Haoran Li, Peng Yuan, Yujia Wang, Youzheng Wu, Xiaodong He, Ying Liu and Bowen Zhou ................................................................. 1

**Extracting Topics with Simultaneous Word Co-occurrence and Semantic Correlation Graphs: Neural Topic Modeling for Short Texts**  
Yiming Wang, Ximing Li, Xiaotang Zhou and Jihong Ouyang ......................... 18

**Self-supervised Contrastive Cross-Modality Representation Learning for Spoken Question Answering**  
Chenyu You, Nuo Chen and Yuexian Zou .................................................. 28

**Language Clustering for Multilingual Named Entity Recognition**  
Kyle Shaffer ............................................................................................... 40

**Neural News Recommendation with Collaborative News Encoding and Structural User Encoding**  
Zhiming Mao, Xingshan Zeng and Kam-Fai Wong ........................................ 46

**Self-Teaching Machines to Read and Comprehend with Large-Scale Multi-Subject Question-Answering Data**  
Dian Yu, Kai Sun, Dong Yu and Claire Cardie ........................................... 56

**A Web Scale Entity Extraction System**  
Xuanting Cai, Quanbin Ma, Jianyu Liu, Pan Li, Qi Zeng, Zhengkan Yang and Pushkar Tripathi .............................................................. 69

**Joint Multimedia Event Extraction from Video and Article**  
Brian Chen, Xudong Lin, Christopher Thomas, Manling Li, Shoya Yoshida, Lovish Chum, Heng Ji and Shih-Fu Chang ........................................................................ 74

**Fine-grained Semantic Alignment Network for Weakly Supervised Temporal Language Grounding**  
Yuechen Wang, Wengang Zhou and Houqiang Li ......................................... 89

**Factual Consistency Evaluation for Text Summarization via Counterfactual Estimation**  
Yuexiang Xie, Fei Sun, Yang Deng, Yaliang Li and Bolin Ding ....................... 100

**Cross-Modal Retrieval Augmentation for Multi-Modal Classification**  
Shir Gur, Natalia Neverova, Chris Stauffer, Ser-Nam Lim, Douwe Kiela and Austin Reiter ........................................................... 111

**HiTRANS: A Hierarchical Transformer Network for Nested Named Entity Recognition**  
Zhiwei Yang, Jing Ma, Hechang Chen, Yunke Zhang and Yi Chang ............. 124

**Improving Embedding-based Large-scale Retrieval via Label Enhancement**  
Peiyang Liu, Xi Wang, Sen Wang, Wei Ye, Xiangyu Xi and Shikun Zhang .......... 133

**Improving Privacy Guarantee and Efficiency of Latent Dirichlet Allocation Model Training Under Differential Privacy**  
Tao Huang and Hong Chen ....................................................................... 143

**Generating Mammography Reports from Multi-view Mammograms with BERT**  
Alexander Yalunin, Elena Sokolova, Ilya Burenko, Alexander Ponomarchuk, Olga Puchkova and Dmitriy Umerenkov ........................................... 153
Euphemistic Phrase Detection by Masked Language Model
Wanzheng Zhu and Suma Bhat ................................................................. 163

Decomposing Complex Questions Makes Multi-Hop QA Easier and More Interpretable
Ruiliu Fu, Han Wang, Xuejun Zhang, Jun Zhou and Yonghong Yan ....................... 169

Segmenting Natural Language Sentences via Lexical Unit Analysis
Yangming Li, Lemao Liu and Shuming Shi .................................................. 181

Dense Hierarchical Retrieval for Open-domain Question Answering
Ye Liu, Kazuma Hashimoto, Yingbo Zhou, Semih Yavuz, Caiming Xiong and Philip Yu .... 188

Visually Grounded Concept Composition
Bowen Zhang, Hexiang Hu, Linlu Qiu, Peter Shaw and Fei Sha ............................. 201

Compositional Networks Enable Systematic Generalization for Grounded Language Understanding
Yen-Ling Kuo, Boris Katz and Andrei Barbu .................................................. 216

An Unsupervised Method for Building Sentence Simplification Corpora in Multiple Languages
Xinyu Lu, Jipeng Qiang, Yun Li, Yunhao Yuan and Yi Zhu .................................. 227

WhiteningBERT: An Easy Unsupervised Sentence Embedding Approach
Junjie Huang, Duyu Tang, Wanjun Zhong, Shuai Lu, Linjun Shou, Ming Gong, Daxin Jiang and Nan Duan .......................................................... 238

TWEETSUMM - A Dialog Summarization Dataset for Customer Service
Guy Feigenblat, Chulaka Gunasekara, Benjamin Sznajder, Sachindra Joshi, David Konopnicki and Ranit Aharonov ................................. 245

Discourse-Based Sentence Splitting
Liam Cripwell, Joël Legrand and Claire Gardent ............................................. 261

Multi-Task Dense Retrieval via Model Uncertainty Fusion for Open-Domain Question Answering
Minghan Li, Ming Li, Kun Xiong and Jimmy Lin ............................................. 274

Mining the Cause of Political Decision-Making from Social Media: A Case Study of COVID-19 Policies across the US States
Zhijing Jin, Zeyu Peng, Tejas Vaidhya, Bernhard Schoelkopf and Rada Mihalcea ........... 288

Self-Attention Graph Residual Convolutional Networks for Event Detection with dependency relations
Anan Liu, Ning Xu and Haozhe Liu ............................................................. 302

Mixup Decoding for Diverse Machine Translation
Jicheng Li, Pengzhi Gao, Xuanfu Wu, Yang Feng, Zhongjun He, Hua Wu and Haifeng Wang . 312

An Alignment-Agnostic Model for Chinese Text Error Correction
Liying Zheng, Yue Deng, Weishun Song, Liang Xu and Jing Xiao ......................... 321

Reasoning Visual Dialog with Sparse Graph Learning and Knowledge Transfer
Gi-Cheon Kang, Junseok Park, Hwaran Lee, Byoung-Tak Zhang and Jin-Hwa Kim .......... 327

Exploring Sentence Community for Document-Level Event Extraction
Yusheng Huang and Weijia Jia ..................................................................... 340
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHOSe Heritage: Classification of UNESCO World Heritage Statements of &quot;Outstanding Universal Value&quot; with Soft Labels</td>
<td>Nan Bai, Renqian Luo, Pirouz Nourian and Ana Pereira Roders</td>
<td>366</td>
</tr>
<tr>
<td>P-INT: A Path-based Interaction Model for Few-shot Knowledge Graph Completion</td>
<td>Jingwen Xu, Jing Zhang, Xirui Ke, Yuxiao Dong, Hong Chen, Cuiping Li and Yongbin Liu</td>
<td>385</td>
</tr>
<tr>
<td>Cartography Active Learning</td>
<td>Mike Zhang and Barbara Plank</td>
<td>395</td>
</tr>
<tr>
<td>GooAQ: Open Question Answering with Diverse Answer Types</td>
<td>Daniel Khashabi, Amos Ng, Tushar Khot, Ashish Sabharwal, Hannaneh Hajishirzi and Chris Callison-Burch</td>
<td>421</td>
</tr>
<tr>
<td>Attention Weights in Transformer NMT Fail Aligning Words Between Sequences but Largely Explain Model Predictions</td>
<td>Javier Ferrando and Marta R. Costa-jussà</td>
<td>434</td>
</tr>
<tr>
<td>BFClass: A Backdoor-free Text Classification Framework</td>
<td>Zichao Li, Dheeraj Mekala, Chengyu Dong and Jingbo Shang</td>
<td>444</td>
</tr>
<tr>
<td>Multilingual Chart-based Constituency Parse Extraction from Pre-trained Language Models</td>
<td>Taeuk Kim, Bowen Li and Sang-goo Lee</td>
<td>454</td>
</tr>
<tr>
<td>Hyperbolic Geometry is Not Necessary: Lightweight Euclidean-Based Models for Low-Dimensional Knowledge Graph Embeddings</td>
<td>Kai Wang, Yu Liu, Dan Lin and Michael Sheng</td>
<td>464</td>
</tr>
<tr>
<td>CascadeBERT: Accelerating Inference of Pre-trained Language Models via Calibrated Complete Models Cascade</td>
<td>Lei Li, Yankai Lin, Deli Chen, Shuhuai Ren, Peng Li, Jie Zhou and Xu Sun</td>
<td>475</td>
</tr>
<tr>
<td>Semi-supervised Relation Extraction via Incremental Meta Self-Training</td>
<td>Xuming Hu, Chenwei Zhang, Fukun Ma, Chenyao Liu, Lijie Wen and Philip S. Yu</td>
<td>487</td>
</tr>
<tr>
<td>Keyphrase Generation with Fine-Grained Evaluation-Guided Reinforcement Learning</td>
<td>Yichao Luo, Yige Xu, Jiacheng Ye, Xipeng Qiu and Qi Zhang</td>
<td>497</td>
</tr>
<tr>
<td>Improving Knowledge Graph Embedding Using Affine Transformations of Entities Corresponding to Each Relation</td>
<td>Jinfa Yang, Yongjie Shi, Xin Tong, Robin Wang, Taiyan Chen and Xianghua Ying</td>
<td>508</td>
</tr>
<tr>
<td>Using Question Answering Rewards to Improve Abstractive Summarization</td>
<td>Chulaka Gunasekara, Guy Feigenblat, Benjamin Sznajder, Ranit Aharonov and Sachindra Joshi</td>
<td>518</td>
</tr>
<tr>
<td>Effect Generation Based on Causal Reasoning</td>
<td>Feiteng Mu, Wenjie Li and Zhipeng Xie</td>
<td>527</td>
</tr>
</tbody>
</table>
Distilling Word Meaning in Context from Pre-trained Language Models
Yuki Arase and Tomoyuki Kajiwara ................................................................. 534

Unseen Entity Handling in Complex Question Answering over Knowledge Base via Language Generation
Xin Huang, Jung-Jae Kim and Bowei Zou .......................................................... 547

Bidirectional Hierarchical Attention Networks based on Document-level Context for Emotion Cause Extraction
Guimin Hu, Guangming Lu and Yi Zhao ............................................................. 558

Distantly Supervised Relation Extraction in Federated Settings
Dianbo Sui, Yubo Chen, Kang Liu and Jun Zhao ............................................... 569

Casting the Same Sentiment Classification Problem
Erik Körner, Ahmad Dawar Hakimi, Gerhard Heyer and Martin Potthast ................ 584

Detecting Compositionally Out-of-Distribution Examples in Semantic Parsing
Denis Lukovnikov, Sina Daubener and Asja Fischer ........................................... 591

Saliency-based Multi-View Mixed Language Training for Zero-shot Cross-lingual Classification
Siyu Lai, Hui Huang, Dong Jing, Yufeng Chen, Jinan Xu and Jian Liu ..................... 599

Fighting the COVID-19 Infodemic: Modeling the Perspective of Journalists, Fact-Checkers, Social Media Platforms, Policy Makers, and the Society

FANATIC: FAst Noise-Aware TopIc Clustering
Ari Silburt, Anja Subasic, Evan Thompson, Carmeline Dsilva and Tarec Fares ............ 650

Stream-level Latency Evaluation for Simultaneous Machine Translation
Javier Iranzo-Sánchez, Jorge Civera Saiz and Alfons Juan ..................................... 664

TSDAE: Using Transformer-based Sequential Denoising Auto-Encoder for Unsupervised Sentence Embedding Learning
Kexin Wang, Nils Reimers and Iryna Gurevych .................................................. 671

How Suitable Are Subword Segmentation Strategies for Translating Non-Concatenative Morphology?
Chantal Amrhein and Rico Sennrich ................................................................. 689

Rethinking Why Intermediate-Task Fine-Tuning Works
Ting-Yun Chang and Chi-Jen Lu .................................................................. 706

Learn Continually, Generalize Rapidly: Lifelong Knowledge Accumulation for Few-shot Learning
Xisen Jin, Bill Yuchen Lin, Mohammad Rostami and Xiang Ren ............................. 714

Efficient Test Time Adapter Ensembling for Low-resource Language Varieties
Xinyi Wang, Yulia Tsvetkov, Sebastian Ruder and Graham Neubig .......................... 730

An Analysis of Euclidean vs. Graph-Based Framing for Bilingual Lexicon Induction from Word Embedding Spaces
Kelly Marchisio, Youngser Park, Ali Saad-Eldin, Anton Alyakin, Kevin Duh, Carey Priebe and Philipp Koehn ................................................................. 738
How to Select One Among All? An Empirical Study Towards the Robustness of Knowledge Distillation in Natural Language Understanding
Tianda Li, Ahmad Rashid, Aref Jafari, Pranav Sharma, Ali Ghodsi and Mehdi Rezagholizadeh

Recommend for a Reason: Unlocking the Power of Unsupervised Aspect-Sentiment Co-Extraction
Zeyu Li, Wei Cheng, Reema Kshetramade, John Houser, Haifeng Chen and Wei Wang

Learning Hard Retrieval Decoder Attention for Transformers
Hongfei Xu, Qiuhui Liu, Josef van Genabith and Deyi Xiong

Recall and Learn: A Memory-augmented Solver for Math Word Problems
Shifeng Huang, Jiawei Wang, Jiao Xu, Da Cao and Ming Yang

An Uncertainty-Aware Encoder for Aspect Detection
Thi-Nhung Nguyen, Kiem-Hieu Nguyen, Young-In Song and Tuan-Dung Cao

Improving Empathetic Response Generation by Recognizing Emotion Cause in Conversations
Jun Gao, Yuhan Liu, Haolin Deng, Wei Wang, Yu Cao, Jiachen Du and Ruifeng Xu

Probing Across Time: What Does RoBERTa Know and When?
Zeyu Liu, Yizhong Wang, Jungo Kasai, Hannaneh Hajishirzi and Noah A. Smith

Knowledge-Guided Paraphrase Identification
Haoyu Wang, Fenglong Ma, Yaqing Wang and Jing Gao

R2-D2: A Modular Baseline for Open-Domain Question Answering
Martin Fajcik, Martin Docekal, Karel Ondrej and Pavel Smrz

What Does Your Smile Mean? Jointly Detecting Multi-Modal Sarcasm and Sentiment Using Quantum Probability
Yaochen Liu, Yazhou Zhang, Qiuchi Li, Benyou Wang and Dwayne Song

Discovering Representation Sprachbund For Multilingual Pre-Training
Yimin Fan, Yaobo Liang, Alexandre Muzio, Hany Hassan, Houqiang Li, Ming Zhou and Nan Duan

Plan-then-Generate: Controlled Data-to-Text Generation via Planning
Yixuan Su, David Vandyke, Sihui Wang, Yimai Fang and Nigel Collier

Few-Shot Table-to-Text Generation with Prototype Memory
Yixuan Su, Zaiqiao Meng, Simon Baker and Nigel Collier

Leveraging Word-Formation Knowledge for Chinese Word Sense Disambiguation
Hua Zheng, Lei Li, Damai Dai, Deli Chen, Tianyu Liu, Xu Sun and Yang Liu

Exploiting Curriculum Learning in Unsupervised Neural Machine Translation
Jinliang Lu and Jiajun Zhang

Robust Fragment-Based Framework for Cross-lingual Sentence Retrieval
Nattapol Trijakwanich, Peerat Limkonchotiwat, Raheem Sarvar, Wannaphong Phatthiyaphaibun, Ekapol Chuangsawanich and Sarana Nutanong

Towards Improving Adversarial Training of NLP Models
Jin Yong Yoo and Yanjun Qi
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>To Protect and To Serve? Analyzing Entity-Centric Framing of Police Violence</td>
<td>Caleb Ziems and Diyi Yang</td>
<td>957</td>
</tr>
<tr>
<td>Calibrate your listeners! Robust communication-based training for pragmatic speakers</td>
<td>Rose Wang, Julia White, Jesse Mu and Noah Goodman</td>
<td>977</td>
</tr>
<tr>
<td>When Retriever-Reader Meets Scenario-Based Multiple-Choice Questions</td>
<td>ZiXian Huang, Ao Wu, Yulin Shen, Gong Cheng and Yuzhong Qu</td>
<td>985</td>
</tr>
<tr>
<td>Structured abbreviation expansion in context</td>
<td>Kyle Gorman, Christo Kirov, Brian Roark and Richard Sproat</td>
<td>995</td>
</tr>
<tr>
<td>Task-adaptive Pre-training and Self-training are Complementary for Natural Language Understanding</td>
<td>Shiyang Li, Semih Yavuz, Wenhu Chen and Xifeng Yan</td>
<td>1006</td>
</tr>
<tr>
<td>CNNBiF: CNN-based Bigram Features for Named Entity Recognition</td>
<td>Chul Sung, Vaibhava Goel, Etienne Marcheret, Steven Rennie and David Nahamoo</td>
<td>1016</td>
</tr>
<tr>
<td>Compositional Generalization via Semantic Tagging</td>
<td>Hao Zheng and Mirella Lapata</td>
<td>1022</td>
</tr>
<tr>
<td>Towards Document-Level Paraphrase Generation with Sentence Rewriting and Reordering</td>
<td>Zhe Lin, Yitao Cai and Xiaojun Wan</td>
<td>1033</td>
</tr>
<tr>
<td>Exploring Decomposition for Table-based Fact Verification</td>
<td>Xiaoyu Yang and Xiaodan Zhu</td>
<td>1045</td>
</tr>
<tr>
<td>Diversity and Consistency: Exploring Visual Question-Answer Pair Generation</td>
<td>Sen Yang, Qingyu Zhou, Dawei Feng, Yang Liu, Chao Li, Yunbo Cao and Dongsheng Li</td>
<td>1053</td>
</tr>
<tr>
<td>Entity-level Cross-modal Learning Improves Multi-modal Machine Translation</td>
<td>Xin Huang, Jiajun Zhang and Chengqing Zong</td>
<td>1067</td>
</tr>
<tr>
<td>Learning to Ground Visual Objects for Visual Dialog</td>
<td>Feilong Chen, Xiuyi Chen, Can Xu and Daxin Jiang</td>
<td>1081</td>
</tr>
<tr>
<td>KERS: A Knowledge-Enhanced Framework for Recommendation Dialog Systems with Multiple Subgoals</td>
<td>Jun Zhang, Yan Yang, Chencai Chen, Liang He and Zhou Yu</td>
<td>1092</td>
</tr>
<tr>
<td>Less Is More: Domain Adaptation with Lottery Ticket for Reading Comprehension</td>
<td>Haichao Zhu, Zekun Wang, Heng Zhang, Ming Liu, Sendong Zhao and Bing Qin</td>
<td>1102</td>
</tr>
<tr>
<td>Effectiveness of Pre-training for Few-shot Intent Classification</td>
<td>Haode Zhang, Yuwei Zhang, Li-Ming Zhan, Jiaxin Chen, Guangyuan Shi, Xiao-Ming Wu and Albert Y.S. Lam</td>
<td>1114</td>
</tr>
<tr>
<td>Improving Abstractive Dialogue Summarization with Hierarchical Pretraining and Topic Segment</td>
<td>MengNan Qi, Hao Liu, YuZhuo Fu and Ting Liu</td>
<td>1121</td>
</tr>
<tr>
<td>Learning to Answer Psychological Questionnaire for Personality Detection</td>
<td>Feifan Yang, Tao Yang, Xiaojun Quan and Qinliang Su</td>
<td>1131</td>
</tr>
<tr>
<td>Exploiting Reasoning Chains for Multi-hop Science Question Answering</td>
<td>Weiwen Xu, Yang Deng, Huihui Zhang, Deng Cai and Wai Lam</td>
<td>1143</td>
</tr>
</tbody>
</table>
Winnowing Knowledge for Multi-choice Question Answering
Yeqiu Li, Bowei Zou, Zhifeng Li, Ai Ti Aw, Yu Hong and Qiaoming Zhu

Neural Media Bias Detection Using Distant Supervision With BABE - Bias Annotations By Experts
Timo Spinde, Manuel Plank, Jan-David Krieger, Terry Ruas, Bela Gipp and Akiko Aizawa

Learning and Evaluating a Differentially Private Pre-trained Language Model
Shlomo Hoory, Amir Feder, Avichai Tendler, Sofia Erell, Alon Peled-Cohen, Itay Laish, Hootan Nakhost, Uri Stemmer, Ayelet Benjamini, Avinatan Hassidim and Yossi Matias

Simulated Chats for Building Dialog Systems: Learning to Generate Conversations from Instructions
Biswaesh Mohapatra, Gaurav Pandey, Danish Contractor and Sachindra Joshi

Past, Present, and Future: Conversational Emotion Recognition through Structural Modeling of Psychological Knowledge
Jiangnan Li, Zheng Lin, Peng Fu and Weiping Wang

An unsupervised framework for tracing textual sources of moral change
Aida Ramezani, Zining Zhu, Frank Rudzicz and Yang Xu

Topic-Aware Contrastive Learning for Abstractive Dialogue Summarization
Junpeng Liu, Yanyan Zou, Hainan Zhang, Hongshen Chen, Zhuoye Ding, Caixia Yuan and Xiaojie Wang

TWT: Table with Written Text for Controlled Data-to-Text Generation
Tongliang Li, Lei Fang, Jian-Guang Lou and Zhoujun Li

ArabicTransformer: Efficient Large Arabic Language Model with Funnel Transformer and ELECTRA Objective
Sultan Alrowili and Vijay Shanker

Which is Making the Contribution: Modulating Unimodal and Cross-modal Dynamics for Multimodal Sentiment Analysis
Ying Zeng, Sijie Mai and Haifeng Hu

CVAE-based Re-anchoring for Implicit Discourse Relation Classification
Zujun Dou, Yu Hong, Yu Sun and Guodong Zhou

Combining Curriculum Learning and Knowledge Distillation for Dialogue Generation
Qingqing Zhu, Xiuying Chen, Pengfei Wu, JunFei Liu and Dongyan Zhao

Improving End-to-End Task-Oriented Dialog System with A Simple Auxiliary Task
Yohan Lee

EDTC: A Corpus for Discourse-Level Topic Chain Parsing
Longyin Zhang, Xin Tan, Fang Kong and Guodong Zhou

Multilingual Neural Machine Translation: Can Linguistic Hierarchies Help?
Fahimeh Saleh, Wray Buntine, Gholamreza Haffari and Lan Du

Self Question-answering: Aspect-based Sentiment Analysis by Role Flipped Machine Reading Comprehension
Guoxin Yu, Jiwei Li, Ling Luo, Yuxian Meng, Xiang Ao and Qing He

Generalization in Text-based Games via Hierarchical Reinforcement Learning
Yunqiu Xu, Meng Fang, Ling Chen, Yali Du and Chengqi Zhang
MiRANews: Dataset and Benchmarks for Multi-Resource-Assisted News Summarization
Xinnuo Xu, Ondřej Dušek, Shashi Narayan, Verena Rieser and Ioannis Konstas ................. 1541

A Conditional Generative Matching Model for Multi-lingual Reply Suggestion
Budhaditya Deb, Guoqing Zheng, Milad Shokouhi and Ahmed Hassan Awadallah .......... 1553

Rethinking Sentiment Style Transfer
Ping Yu, Yang Zhao, Chunyuan Li and Changyou Chen ........................................... 1569

HypoGen: Hyperbole Generation with Commonsense and Counterfactual Knowledge
Yufei Tian, Arvind krishna Sridhar and Nanyun Peng .............................................. 1583

Profiling News Discourse Structure Using Explicit Subtopic Structures Guided Critics
Prafulla Kumar Choubey and Ruihong Huang ........................................................... 1594

ProtoInfoMax: Prototypical Networks with Mutual Information Maximization for Out-of-Domain Detection
Iftitahu Nimah, Meng Fang, Vlado Menkovski and Mykola Pechenizkiy ....................... 1606

Learning from Language Description: Low-shot Named Entity Recognition via Decomposed Framework
Yaqing Wang, Haoda Chu, Chao Zhang and Jing Gao ................................................. 1618

BERT might be Overkill: A Tiny but Effective Biomedical Entity Linker based on Residual Convolutional Neural Networks
Tuan Lai, Heng Ji and ChengXiang Zhai ................................................................. 1631

Char2Subword: Extending the Subword Embedding Space Using Robust Character Compositionality
Gustavo Aguilar, Bryan McCann, Tong Niu, Nazneen Rajani, Nitish Shirish Keskar and Thamar Solorio ................................................................. 1640

Exploring Multitask Learning for Low-Resource Abstractive Summarization
Ahmed Magooda, Diane Litman and Mohamed Elaraby .............................................. 1652

Conical Classification For Efficient One-Class Topic Determination
Sameer Khanna ........................................................................................................... 1662

Improving Dialogue State Tracking with Turn-based Loss Function and Sequential Data Augmentation
Jarana Manotumruksa, Jeff Dalton, Edgar Meij and Emine Yilmaz ............................ 1674

TIAGE: A Benchmark for Topic-Shift Aware Dialog Modeling
Huiyuan Xie, Zhenghao Liu, Chenyan Xiong, Zhiyuan Liu and Ann Copestake .......... 1684

Optimal Neural Program Synthesis from Multimodal Specifications
Xi Ye, Qiaochu Chen, Isil Dillig and Greg Durrett ....................................................... 1691

Sent2Span: Span Detection for PICO Extraction in the Biomedical Text without Span Annotations
Shifeng Liu, Yifang Sun, Bing Li, Wei Wang, Florence T. Bourgeois and Adam G. Dunn .... 1705

When in Doubt: Improving Classification Performance with Alternating Normalization
Menglin Jia, Austin Reiter, Ser-Nam Lim, Yoav Artzi and Claire Cardie ...................... 1716

APGN: Adversarial and Parameter Generation Networks for Multi-Source Cross-Domain Dependency Parsing
Ying Li, Meishan Zhang, Zhenghua Li, Min Zhang, Zhefeng Wang, baoxing Huai and Nicholas Jing Yuan ................................................................. 1724
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Let Your Characters Tell Their Story&quot;: A Dataset for Character-Centric Narrative Understanding</td>
<td>Faeze Brahman, Meng Huang, Oyvind Tafjord, Chao Zhao, Mrinmaya Sachan and Snigdha Chaturvedi</td>
<td>1734</td>
</tr>
<tr>
<td>Towards Developing a Multilingual and Code-Mixed Visual Question Answering System by Knowledge Distillation</td>
<td>Humair Raj Khan, Deepak Gupta and Asif Ekbal</td>
<td>1753</td>
</tr>
<tr>
<td>An Iterative Multi-Knowledge Transfer Network for Aspect-Based Sentiment Analysis</td>
<td>Yunlong Liang, Fandong Meng, Jinchao Zhang, Yufeng Chen, Jinan Xu and Jie Zhou</td>
<td>1768</td>
</tr>
<tr>
<td>Semantic Alignment with Calibrated Similarity for Multilingual Sentence Embedding</td>
<td>Jiyeon Ham and Eun-Sol Kim</td>
<td>1781</td>
</tr>
<tr>
<td>fBERT: A Neural Transformer for Identifying Offensive Content</td>
<td>Diptanu Sarkar, Marcos Zampieri, Tharindu Ranasinghe and Alexander Ororbia</td>
<td>1792</td>
</tr>
<tr>
<td>WIKIBIAS: Detecting Multi-Span Subjective Biases in Language</td>
<td>Yang Zhong, Jingfeng Yang, Wei Xu and Diyi Yang</td>
<td>1799</td>
</tr>
<tr>
<td>UncIE: Explicitly Leveraging Semantic Similarity to Reduce the Parameters of Word Embeddings</td>
<td>Zhi Li, Yuchen Zhai, Chengu Chen, Minghui Qiu, Kailiang Li and Yin Zhang</td>
<td>1815</td>
</tr>
<tr>
<td>Grounded Graph Decoding improves Compositional Generalization in Question Answering</td>
<td>Yu Gai, Paras Jain, Wendi Zhang, Joseph Gonzalez, Dawn Song and Ion Stoica</td>
<td>1829</td>
</tr>
<tr>
<td>Enhancing Visual Dialog Questioner with Entity-based Strategy Learning and Augmented Guesser</td>
<td>Duo Zheng, Zipeng Xu, Fandong Meng, Xiaojie Wang, Jiaan Wang and Jie Zhou</td>
<td>1839</td>
</tr>
<tr>
<td>A Pretraining Numerical Reasoning Model for Ordinal Constrained Question Answering on Knowledge Base</td>
<td>Yu Feng, Jing Zhang, Gaole He, Wayne Xin Zhao, Lemao Liu, Quan Liu, Cuiping Li and Hong Chen</td>
<td>1852</td>
</tr>
<tr>
<td>RoR: Read-over-Read for Long Document Machine Reading Comprehension</td>
<td>Jing Zhao, Junwei Bao, Yifan Wang, Yongwei Zhou, Youzheng Wu, Xiaodong He and Bowen Zhou</td>
<td>1862</td>
</tr>
<tr>
<td>Language Resource Efficient Learning for Captioning</td>
<td>Jia Chen, Yike Wu, Shiwan Zhao and Qin Jin</td>
<td>1887</td>
</tr>
<tr>
<td>Translation as Cross-Domain Knowledge: Attention Augmentation for Unsupervised Cross-Domain Segmenting and Labeling Tasks</td>
<td>Ruixuan Luo, Yi Zhang, Sishuo Chen and Xu Sun</td>
<td>1896</td>
</tr>
<tr>
<td>ContractNLI: A Dataset for Document-level Natural Language Inference for Contracts</td>
<td>Yuta Koreeda and Christopher Manning</td>
<td>1907</td>
</tr>
<tr>
<td>Japanese Zero Anaphora Resolution Can Benefit from Parallel Texts Through Neural Transfer Learning</td>
<td>Masato Umakoshi, Yugo Murawaki and Sadao Kurohashi</td>
<td>1920</td>
</tr>
<tr>
<td>Title</td>
<td>Authors</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Grouped-Attention for Content-Selection and Content-Plan Generation</td>
<td>Bayu Distiawan Trisedya, Xiaojie Wang, Jianzhong Qi, Rui Zhang and Qingjun Cui</td>
<td>1935</td>
</tr>
<tr>
<td>An Explicit-Joint and Supervised-Contrastive Learning Framework for Few-Shot Intent Classification and Slot Filling</td>
<td>Han Liu, Feng Zhang, Xiaotong Zhang, Siyang Zhao and Xianchao Zhang</td>
<td>1945</td>
</tr>
<tr>
<td>Retrieve, Discriminate and Rewrite: A Simple and Effective Framework for Obtaining Affective Response in Retrieval-Based Chatbots</td>
<td>Xin Lu, Yijian Tian, Yanyan Zhao and Bing Qin</td>
<td>1956</td>
</tr>
<tr>
<td>Span Fine-tuning for Pre-trained Language Models</td>
<td>Rongzhou Bao, Zhuosheng Zhang and Hai Zhao</td>
<td>1970</td>
</tr>
<tr>
<td>DIRECT: Direct and Indirect Responses in Conversational Text Corpus</td>
<td>Junya Takayama, Tomoyuki Kajiwara and Yuki Arase</td>
<td>1980</td>
</tr>
<tr>
<td>Retrieval, Analogy, and Composition: A framework for Compositional Generalization in Image Captioning</td>
<td>Zhan Shi, Hui Liu, Martin Renqiang Min, Christopher Malon, Li Erran Li and Xiaodan Zhu</td>
<td>1990</td>
</tr>
<tr>
<td>TURINGBENCH: A Benchmark Environment for Turing Test in the Age of Neural Text Generation</td>
<td>Adaku Uchendu, Zeyu Ma, Thai Le, Rui Zhang and Dongwon Lee</td>
<td>2001</td>
</tr>
<tr>
<td>Say ‘YES’ to Positivity: Detecting Toxic Language in Workplace Communications</td>
<td>Meghana Moorthy Bhat, saghar Hosseini, Ahmed Hassan Awadallah, Paul Bennett and Weisheng Li</td>
<td>2017</td>
</tr>
<tr>
<td>Natural SQL: Making SQL Easier to Infer from Natural Language Specifications</td>
<td>Yujian Gan, Xinyun Chen, Jinxia Xie, Matthew Purver, John R. Woodward, John Drake and Qiaofu Zhang</td>
<td>2030</td>
</tr>
<tr>
<td>Mitigating Data Scarceness through Data Synthesis, Augmentation and Curriculum for Abstractive Summarization</td>
<td>Ahmed Magooda and Diane Litman</td>
<td>2043</td>
</tr>
<tr>
<td>Self- and Pseudo-self-supervised Prediction of Speaker and Key-utterance for Multi-party Dialogue Reading Comprehension</td>
<td>Yiyang Li and Hai Zhao</td>
<td>2053</td>
</tr>
<tr>
<td>Few-Shot Novel Concept Learning for Semantic Parsing</td>
<td>Soham Dan, Osbert Bastani and Dan Roth</td>
<td>2064</td>
</tr>
<tr>
<td>Compositional Data and Task Augmentation for Instruction Following</td>
<td>Soham Dan, Xinran Han and Dan Roth</td>
<td>2076</td>
</tr>
<tr>
<td>Are Factuality Checkers Reliable? Adversarial Meta-evaluation of Factuality in Summarization</td>
<td>Yiran Chen, Pengfei Liu and Xipeng Qiu</td>
<td>2082</td>
</tr>
<tr>
<td>On the Effects of Transformer Size on In- and Out-of-Domain Calibration</td>
<td>Soham Dan and Dan Roth</td>
<td>2096</td>
</tr>
<tr>
<td>Detecting Polarized Topics Using Partisanship-aware Contextualized Topic Embeddings</td>
<td>Zihao He, Negar Mokhberian, António Câmara, Andres Abeliuk and Kristina Lerman</td>
<td>2102</td>
</tr>
</tbody>
</table>
GenerativeRE: Incorporating a Novel Copy Mechanism and Pretrained Model for Joint Entity and Relation Extraction
Jiarun Cao and Sophia Ananiadou .......................................................... 2119

Re-entry Prediction for Online Conversations via Self-Supervised Learning
Lingzhi Wang, Xingshan Zeng, Huang Hu, Kam-Fai Wong and Daxin Jiang ............... 2127

proScript: Partially Ordered Scripts Generation
Keisuke Sakaguchi, Chandra Bhagavatula, Ronan Le Bras, Niket Tandon, Peter Clark and Yejin Choi .......................................................... 2138

Speaker Turn Modeling for Dialogue Act Classification
Zihao He, Leili Tavabi, Kristina Lerman and Mohammad Soleymani ...................... 2150

Unsupervised Domain Adaptation Method with Semantic-Structural Alignment for Dependency Parsing
Boda Lin, Mingzheng Li, Si Li and Yong Luo .............................................. 2158

Devil’s Advocate: Novel Boosting Ensemble Method from Psychological Findings for Text Classification
Hwiyeol Jo, Jaeseo Lim and Byoung-Tak Zhang ......................................... 2168

SideControl: Controlled Open-domain Dialogue Generation via Additive Side Networks
Wanyu Du and Yangfeng Ji ................................................................. 2175

Is BERT a Cross-Disciplinary Knowledge Learner? A Surprising Finding of Pre-trained Models’ Transferability
Wei-Tsung Kao and Hung-ji Lee .......................................................... 2195

Geo-BERT Pre-training Model for Query Rewriting in POI Search
Xiao Liu, Juan Hu, Qi Shen and Huan Chen .............................................. 2209

Leveraging Bidding Graphs for Advertiser-Aware Relevance Modeling in Sponsored Search
Shuxian Bi, Chaozhuo Li, Xiao Han, Zheng Liu, Xing Xie, Haizhen Huang and Zengxuan Wen 2215

GPT3Mix: Leveraging Large-scale Language Models for Text Augmentation
Kang Min Yoo, Dongju Park, Jaewook Kang, Sang-Woo Lee and Woomyoung Park ............ 2225

Context-aware Entity Typing in Knowledge Graphs
Weiran Pan, Wei Wei and Xian-Ling Mao .................................................. 2240

Attribute Alignment: Controlling Text Generation from Pre-trained Language Models
Dian Yu, Zhou Yu and Kenji Sagae ....................................................... 2251

Generate & Rank: A Multi-task Framework for Math Word Problems
Jianhao Shen, Yichun Yin, Lin Li, Lifeng Shang, Xin Jiang, Ming Zhang and Qun Liu ...... 2269

MIRTT: Learning Multimodal Interaction Representations from Trilinear Transformers for Visual Question Answering
Junjie Wang, Yatai Ji, Jiaqi Sun, Yujiu Yang and Tetsuya Sakai .......................... 2280

UniteD-SRL: A Unified Dataset for Span- and Dependency-Based Multilingual and Cross-Lingual Semantic Role Labeling
Rocco Tripodi, Simone Conia and Roberto Navigli .................................... 2293
Enhancing Dual-Encoders with Question and Answer Cross-Embeddings for Answer Retrieval
Yanmeng Wang, Jun Bai, Ye Wang, Jianfei Zhang, Wenge Rong, Zongcheng Ji, Shaojun Wang and Jing Xiao ................................................................. 2306

A Neural Graph-based Local Coherence Model
Mohsen Mesgar, Leonardo F. R. Ribeiro and Iryna Gurevych ...................................................... 2316

GiBERT: Enhancing BERT with Linguistic Information using a Lightweight Gated Injection Method
Nicole Peinelt, Marek Rei and Maria Liakata .......................................................... 2322

RollingLDA: An Update Algorithm of Latent Dirichlet Allocation to Construct Consistent Time Series from Textual Data
Jonas Rieger, Carsten Jentsch and Jörg Rahnenführer ..................................................... 2337

What If Sentence-hood is Hard to Define: A Case Study in Chinese Reading Comprehension
Jiawei Wang, Hai Zhao, Yinggong Zhao and Libin Shen ............................................... 2348

Refining BERT Embeddings for Document Hashing via Mutual Information Maximization
Zijing Ou, Qinliang Su, Jianxing Yu, Ruihui Zhao, Yefeng Zheng and Bang Liu ............... 2360

REBEL: Relation Extraction By End-to-end Language generation
Pere-Lluis Huguet Cabot and Roberto Navigli ......................................................................... 2370

Wine is not v i n. On the Compatibility of Tokenizations across Languages
Antonis Maronikolakis, Philipp Dufter and Hinrich Schütze ............................................. 2382

Temporal Adaptation of BERT and Performance on Downstream Document Classification: Insights from Social Media
Paul Röttger and Janet Pierrehumbert .................................................................................. 2400

Skim-Attention: Learning to Focus via Document Layout
Laura Nguyen, Thomas Scialom, Jacopo Staiano and Benjamin Piwowarski .................... 2413

Attention-based Contrastive Learning for Winograd Schemas
Tassilo Klein and Moin Nabi ................................................................................................. 2428

Give the Truth: Incorporate Semantic Slot into Abstractive Dialogue Summarization
Lulu Zhao, Weihao Zeng, Weiran Xu and Jun Guo ............................................................ 2435

Challenges in Detoxifying Language Models
Johannes Welbl, Amelia Glaese, Jonathan Uesato, Sumanth Dathathri, John Mellor, Lisa Anne Hendricks, Kirsty Anderson, Pushmeet Kohli, Ben Coppin and Po-Sen Huang ................. 2447

Collecting a Large-Scale Gender Bias Dataset for Coreference Resolution and Machine Translation
Shahar Levy, Koren Lazar and Gabriel Stanovsky .............................................................. 2470

Competence-based Curriculum Learning for Multilingual Machine Translation
Mingliang Zhang, Fandong Meng, Yunhai Tong and Jie Zhou ........................................... 2481

Informed Sampling for Diversity in Concept-to-Text NLG
Giulio Zhou and Gerasimos Lampouras ............................................................................. 2494

Novel Natural Language Summarization of Program Code via Leveraging Multiple Input Representations
Fuxiang Chen, Mijung Kim and Jaegul Choo ................................................................. 2510
WikiNEuRal: Combined Neural and Knowledge-based Silver Data Creation for Multilingual NER
Simone Tedeschi, Valentino Maiorca, Niccolò Campolungo, Francesco Ceconi and Roberto Navigli ................................................................. 2521

Beyond Grammatical Error Correction: Improving L1-influenced research writing in English using pre-trained encoder-decoder models
Gustavo Zomer and Ana Frankenberg-Garcia ................................................................. 2534

Classification and Geotemporal Analysis of Quality-of-Life Issues in Tenant Reviews
Adam Haber and Zeev Waks ................................................................. 2541

Probing Pre-trained Language Models for Semantic Attributes and their Values
Meriem Beloucif and Chris Biemann ................................................................. 2554

Uncovering the Limits of Text-based Emotion Detection
Nurudin Alvarez-Gonzalez, Andreas Kaltenbrunner and Vicenç Gómez ................................................................. 2560

Named Entity Recognition for Entity Linking: What Works and What’s Next
Simone Tedeschi, Simone Conia, Francesco Ceconi and Roberto Navigli ................................................................. 2584

Learning Numeracy: A Simple Yet Effective Number Embedding Approach Using Knowledge Graph
Hanyu Duan, Yi Yang and Kar Yan Tam ................................................................. 2597

Weakly Supervised Semantic Parsing by Learning from Mistakes
Jiaqi Guo, Jian-Guang Lou, Ting Liu and Dongmei Zhang ................................................................. 2603

CodeQA: A Question Answering Dataset for Source Code Comprehension
Chenxiao Liu and Xiaojun Wan ................................................................. 2618

Subword Mapping and Anchoring across Languages
Giorgos Vernikos and Andrei Popescu-Belis ................................................................. 2633

CDLM: Cross-Document Language Modeling
Avi Caciularu, Arman Cohan, Iz Beltagy, Matthew Peters, Arie Cattan and Ido Dagan ................................................................. 2648

Patterns of Polysemy and Homonymy in Contextualised Language Models
Janosch Haber and Massimo Poesio ................................................................. 2663

Cross-Lingual Leveled Reading Based on Language-Invariant Features
Simin Rao, Hua Zheng and Sujian Li ................................................................. 2677

Controlled Neural Sentence-Level Reframing of News Articles
Wei-Fan Chen, Khalid Al Khatib, Benno Stein and Henning Wachsmuth ................................................................. 2683

DialogueTRM: Exploring Multi-Modal Emotional Dynamics in a Conversation
Yuzhao Mao, Guang Liu, Xiaojie Wang, Weiguo Gao and Xuan Li ................................................................. 2694

Adversarial Examples for Evaluating Math Word Problem Solvers
Vivek Kumar, Rishabh Maheshwary and Vikram Pudi ................................................................. 2705

Improving Numerical Reasoning Skills in the Modular Approach for Complex Question Answering on Text
Xiao-Yu Guo, Yuan-Fang Li and Gholamreza Haffari ................................................................. 2713

Retrieval Augmented Code Generation and Summarization
Md Rizwan Parvez, Wasi Ahmad, Saikat Chakraborty, Baishakhi Ray and Kai-Wei Chang ................................................................. 2719
Multilingual Translation via Grafting Pre-trained Language Models
Zewei Sun, Mingxuan Wang and Lei Li ................................................................. 2735
AEDA: An Easier Data Augmentation Technique for Text Classification
Akbar Karimi, Leonardo Rossi and Andrea Prati ............................................. 2748
A Comprehensive Comparison of Word Embeddings in Event & Entity Coreference Resolution.
Judicael Poumay and Ashwin Ittoo ............................................................... 2755
Wav-BERT: Cooperative Acoustic and Linguistic Representation Learning for Low-Resource Speech Recognition
Guolin Zheng, Yubei Xiao, Ke Gong, Pan Zhou, Xiaodan Liang and Liang Lin ........ 2765
Multilingual AMR Parsing with Noisy Knowledge Distillation
Deng Cai, Xin Li, Jackie Chun-Sing Ho, Lidong Bing and Wai Lam ..................... 2778
Open-Domain Contextual Link Prediction and its Complementarity with Entailment Graphs
Mohammad Javad Hosseini, Shay B. Cohen, Mark Johnson and Mark Steedman .... 2790
Analysis of Language Change in Collaborative Instruction Following
Anna Effenberger, Rhia Singh, Eva Yan, Alane Suhr and Yoav Artzi ...................... 2803
Counter-Interference Adapter for Multilingual Machine Translation
Yaoming Zhu, Jiangtao Feng, Chengqi Zhao, Mingxuan Wang and Lei Li .............. 2812
Progressive Transformer-Based Generation of Radiology Reports
Farhad Nooralahzadeh, Nicolas Perez Gonzalez, Thomas Frauenfelder, Koji Fujimoto and Michael Krauthammer ................................................................. 2824
"Be nice to your wife! The restaurants are closed": Can Gender Stereotype Detection Improve Sexism Classification?
Patricia Chiril, Farah Benamara and Véronique Moriceau .................................. 2833
Automatic Discrimination between Inherited and Borrowed Latin Words in Romance Languages
Alina Maria Cristea, Liviu P. Dinu, Simona Georgescu, Mihnea-Lucian Mihai and Ana Sabina Uban 2845
Adapting Language Models for Zero-shot Learning by Meta-tuning on Dataset and Prompt Collections
Ruiqi Zhong, Kristy Lee, Zheng Zhang and Dan Klein ....................................... 2856
Knowledge-Interactive Network with Sentiment Polarity Intensity-Aware Multi-Task Learning for Emotion Recognition in Conversations
Yunhe Xie, Kailai Yang, Chengjie Sun, Bingquan Liu and Zhenzhou Ji ............... 2879
Minimizing Annotation Effort via Max-Volume Spectral Sampling
Ariadna Quattoni and Xavier Carreras ............................................................ 2890
On the Complementarity between Pre-Training and Back-Translation for Neural Machine Translation
Xuebo Liu, Longyue Wang, Derek F. Wong, Liang Ding, Lidia S. Chao, Shuming Shi and Zhaopeng Tu ................................................................. 2900
Lexicon-Based Graph Convolutional Network for Chinese Word Segmentation
Kaiyu Huang, Hao Yu, Junpeng Liu, Wei Liu, Jingxiang Cao and Degen Huang .......... 2908
KFCNet: Knowledge Filtering and Contrastive Learning for Generative Commonsense Reasoning
Haonan Li, Yeyun Gong, Jian Jiao, Ruofei Zhang, Timothy Baldwin and Nan Duan .... 2918
Monolingual and Cross-Lingual Acceptability Judgments with the Italian CoLA corpus
Daniela Trotta, Raffaele Guarasci, Elisa Leonardelli and Sara Tonelli ................................. 2929

Hyperbolic Hierarchy-Aware Knowledge Graph Embedding for Link Prediction
Zhe Pan and Peng Wang ................................................................. 2941

A Discourse-Aware Graph Neural Network for Emotion Recognition in Multi-Party Conversation
Yang Sun, Nan Yu and Guohong Fu .............................................. 2949

MeLT: Message-Level Transformer with Masked Document Representations as Pre-Training for Stance Detection
Matthew Matero, Nikita Soni, Niranjan Balasubramanian and H. Andrew Schwartz ............. 2959

LMSOC: An Approach for Socially Sensitive Pretraining
Vivek Kulkarni, Shubhanshu Mishra and Aria Haghighi .................................................. 2967

Extract, Integrate, Compete: Towards Verification Style Reading Comprehension
Chen Zhang, Yuxuan Lai, Yansong Feng and Dongyan Zhao ........................................... 2976

Comparing learnability of two dependency schemes: ‘semantic’ (UD) and ‘syntactic’ (SUD)
Ryszard Tuora, Adam Przepiórkowski and Aleksander Leczkowski .................................. 2987

Argumentation-Driven Evidence Association in Criminal Cases
Yefei Teng and WenHan Chao ....................................................... 2997

Eliminating Sentiment Bias for Aspect-Level Sentiment Classification with Unsupervised Opinion Extraction
Bo Wang, Tao Shen, Guodong Long, Tianyi Zhou and Yi Chang ..................................... 3002

Data Efficient Masked Language Modeling for Vision and Language
Yonatan Bitton, Michael Elhadad, Gabriel Stanovsky and Roy Schwartz .......................... 3013

Improving Multilingual Neural Machine Translation with Auxiliary Source Languages
Weijia Xu, Yuwei Yin, Shuming Ma, Dongdong Zhang and Haoyang Huang ....................... 3029

How Does Fine-tuning Affect the Geometry of Embedding Space: A Case Study on Isotropy
Sara Rajaee and Mohammad Taher Pilehvar ......................................................... 3042

Locality Preserving Sentence Encoding
Changrong Min, Yonghe Chu, Liang Yang, Bo Xu and Hongfei Lin ................................. 3050

Knowledge Representation Learning with Contrastive Completion Coding
Bo Ouyang, Wenbing Huang, Runfa Chen, Zhixing Tan, Yang Liu, Maosong Sun and Jihong Zhu 3061

Knowledge-Enhanced Evidence Retrieval for Counterargument Generation
Yohan Jo, Haneul Yoo, JinYeong Bak, Alice Oh, Chris Reed and Eduard Hovy .................... 3074

Investigating Numeracy Learning Ability of a Text-to-Text Transfer Model
Kuntal Kumar Pal and Chitta Baral ..................................................... 3095

Modeling Mathematical Notation Semantics in Academic Papers
Hwiyeol Jo, Dongyeop Kang, Andrew Head and Marti A. Hearst ................................. 3102
Unpacking the Interdependent Systems of Discrimination: Ableist Bias in NLP Systems through an Intersectonal Lens
Saad Hassan, Matt Huenerfauth and Cecilia Ovesdotter Alm .......................... 3116

Constructing Emotional Consensus and Utilizing Unpaired Data for Empathetic Dialogue Generation
Lei Shen, Jinchao Zhang, Jiao Ou, Xiaofang Zhao and Jie Zhou .......................... 3124

Automatic rule generation for time expression normalization
Wentao Ding, Jianhao Chen, Jinmiao Li and Yuzhong Qu .......................... 3135

RW-KD: Sample-wise Loss Terms Re-Weighting for Knowledge Distillation
Peng Lu, Abbas Ghaddar, Ahmad Rashid, Mehdi Rezagholizadeh, Ali Ghodsi and Philippe Langlais 3145

Visual Cues and Error Correction for Translation Robustness
Zhenhao Li, Marek Rei and Lucia Specia .................................................. 3153

Beyond the Tip of the Iceberg: Assessing Coherence of Text Classifiers
Shane Storks and Joyce Chai .................................................. 3169

Does Pretraining for Summarization Require Knowledge Transfer?
Kundan Krishna, Jeffrey Bigham and Zachary C. Lipton .................................. 3178

Bandits Don’t Follow Rules: Balancing Multi-Facet Machine Translation with Multi-Armed Bandits
Julia Kreutzer, David Vilar and Artem Sokolov .................................. 3190

Sometimes We Want Ungrammatical Translations
Prasanna Parthasarathi, Koustuv Sinha, Joelle Pineau and Adina Williams .......... 3205

An animated picture says at least a thousand words: Selecting Gif-based Replies in Multimodal Dialog
Xingyao Wang and David Jurgens .................................................. 3228

SciCap: Generating Captions for Scientific Figures
Ting-Yao Hsu, C Lee Giles and Ting-Hao Huang .................................. 3258

SentNoB: A Dataset for Analysing Sentiment on Noisy Bangla Texts
Khondoker Ittehadul Islam, Sudipta Kar, Md Saiful Islam and Mohammad Ruhul Amin .... 3265

Translate & Fill: Improving Zero-Shot Multilingual Semantic Parsing with Synthetic Data
Massimo Nicosia, Zhongdi Qu and Yasemin Altun .................................. 3272

NewsBERT: Distilling Pre-trained Language Model for Intelligent News Application
Chuhan Wu, Fangzhao Wu, Yang Yu, Tao Qi, Yongfeng Huang and Qi Liu ............. 3285

SD-QA: Spoken Dialectal Question Answering for the Real World
Fahim Faisal, Sharlina Keshava, Md Mahfuz Ibn Alam and Antonios Anastasopoulos 3296

Orevaoghene Ahia, Julia Kreutzer and Sara Hooker .................................. 3316

Transformer over Pre-trained Transformer for Neural Text Segmentation with Enhanced Topic Coherence
Kelvin Lo, Yuan Jin, Weicong Tan, Ming Liu, Lan Du and Wray Buntine .............. 3334

Self-Supervised Neural Topic Modeling
Seyed Ali Bahrainian, Martin Jaggi and Carsten Eickhoff .......................... 3341
Coreference-aware Surprisal Predicts Brain Response
Evan Jaffe, Byung-Doh Oh and William Schuler 3351

Distilling the Knowledge of Large-scale Generative Models into Retrieval Models for Efficient Open-domain Conversation
Beomsu Kim, Seokjun Seo, Seungju Han, Enkhbayar Erdenee and Buru Chang 3357

Modeling Users and Online Communities for Abuse Detection: A Position on Ethics and Explainability
Pushkar Mishra, Helen Yannakoudakis and Ekaterina Shutova 3374

Detecting Community Sensitive Norm Violations in Online Conversations
Chan Young Park, Julia Mendelsohn, Karthik Radhakrishnan, Kinjal Jain, Tushar Kanakagiri, David Jurgens and Yulia Tsvetkov 3386

SupCL-Seq: Supervised Contrastive Learning for Downstream Optimized Sequence Representations
Hooman Sedghamiz, Shivam Raval, Enrico Santus, Tuka Alhanai and Mohammad Ghassemi 3398

mDAPT: Multilingual Domain Adaptive Pretraining in a Single Model
Rasmus Kær Jørgensen, Mareike Hartmann, Xiang Dai and Desmond Elliott 3404

COSMIC: A Coherence-Aware Generation Metric for Image Descriptions
Mert Inan, Piyush Sharma, Baber Khalid, Radu Soricut, Matthew Stone and Malihe Alikhani 3419

Relation-Guided Pre-Training for Open-Domain Question Answering
Ziniu Hu, Yizhou Sun and Kai-Wei Chang 3431

MURAL: Multimodal, Multitask Representations Across Languages
Aashi Jain, Mandy Guo, Krishna Srinivasan, Ting Chen, Sneha Kudugunta, Chao Jia, Yinfei Yang and Jason Baldridge 3449

AStitchInLanguageModels: Dataset and Methods for the Exploration of Idiomaticity in Pre-Trained Language Models
Harish Tayyar Madabushi, Edward Gow-Smith, Carolina Scarton and Aline Villavicencio 3464

Refine and Imitate: Reducing Repetition and Inconsistency in Persuasion Dialogues via Reinforcement Learning and Human Demonstration
Weiyan Shi, Yu Li, Saurav Sahay and Zhou Yu 3478

A Computational Exploration of Pejorative Language in Social Media
Liviu P. Dinu, Ioan-Bogdan Iordache, Ana Sabina Uban and Marcos Zampieri 3493

Evidence-based Fact-Checking of Health-related Claims
Mourad Sarrouti, Asma Ben Abacha, Yassine Mrabet and Dina Demner-Fushman 3499

Learning and Analyzing Generation Order for Undirected Sequence Models
Yichen Jiang and Mohit Bansal 3513

Automatic Bilingual Markup Transfer
Thomas Zenkel, Joern Wuebker and John DeNero 3524

Exploring a Unified Sequence-To-Sequence Transformer for Medical Product Safety Monitoring in Social Media
Shivam Raval, Hooman Sedghamiz, Enrico Santus, Tuka Alhanai, Mohammad Ghassemi and Emmanuele Chersoni 3534
Disentangling Generative Factors in Natural Language with Discrete Variational Autoencoders
Giangiacomo Mercatali and André Freitas ................................................................. 3547

MSD: Saliency-aware Knowledge Distillation for Multimodal Understanding
Woojeong Jin, Maziar Sanjabi, Shaoliang Nie, Liang Tan, Xiang Ren and Hamed Firooz .... 3557

Do UD Trees Match Mention Spans in Coreference Annotations?
Martin Popel, Zdeněk Žabokrtský, Anna Nedoluzhko, Michal Novák and Daniel Zeman .... 3570

Beyond Distillation: Task-level Mixture-of-Experts for Efficient Inference
Sneha Kudugunta, Yanping Huang, Ankur Bapna, Maxim Krikun, Dmitry Lepikhin, Minh-Thang Luong and Orhan Firat............................................................ 3577

TAG: Gradient Attack on Transformer-based Language Models
Jieren Deng, Yijue Wang, Ji Li, Chenghong Wang, Chao Shang, Hang Liu, Sanguthevar Rajasekaran and Caiwen Ding....................................................... 3600

Generating Realistic Natural Language Counterfactuals
Marcel Robeer, Floris Bex and Ad Feelders ................................................................. 3611

Unsupervised Chunking as Syntactic Structure Induction with a Knowledge-Transfer Approach
Anup Anand Deshmukh, Qianqiu Zhang, Ming Li, Jimmy Lin and Lili Mou ................. 3626

Model-based analysis of brain activity reveals the hierarchy of language in 305 subjects
Charlotte Caucheteux, Alexandre Gramfort and Jean-Remi King ............................. 3635

Gated Transformer for Robust De-noised Sequence-to-Sequence Modelling
Ayan Sengupta, Amit Kumar, Sourabh Kumar Bhattacharjee and Suman Roy ............... 3645

Token-wise Curriculum Learning for Neural Machine Translation
Chen Liang, Haoming Jiang, Xiaodong Liu, Pengcheng He, Weizhu Chen, Jianfeng Gao and Tuo Zhao .......................................................... 3658

RelDiff: Enriching Knowledge Graph Relation Representations for Sensitivity Classification
Hitarth Narvala, Graham McDonald and Iadh Ounis .............................................. 3671

Post-Editing Extractive Summaries by Definiteness Prediction
Jad Kabbara and Jackie Chi Kit Cheung ................................................................. 3682

Leveraging Pretrained Models for Automatic Summarization of Doctor-Patient Conversations
Longxiang Zhang, Renato Negrinho, Arindam Ghosh, Vasudevan Jagannathan, Hamid Reza Hassanzadeh, Thomas Schaaf and Matthew R. Gormley ........................................ 3693

Distilling Knowledge for Empathy Detection
Mahshid Hosseini and Cornelia Caragea .............................................................. 3713

Adapting Entities across Languages and Cultures
Denis Peskov, Viktor Hangya, Jordan Boyd-Graber and Alexander Fraser ................. 3725

ODIST: Open World Classification via Distributionally Shifted Instances
Lei Shu, Yassine Benajiba, Saab Mansour and Yi Zhang ........................................ 3751

LAMAD: A Linguistic Attentional Model for Arabic Text Diacritization
Raeed Al-Sabri and Jianliang Gao ........................................................................... 3757
Sequence-to-Lattice Models for Fast Translation
Yuntian Deng and Alexander Rush ................................................................. 3765

Towards Realistic Single-Task Continuous Learning Research for NER
Justin Payan, Yuval Merhav, He Xie, Satyapriya Krishna, Anil Ramakrishna, Mukund Sridhar and Rahul Gupta ................................................................. 3773

Retrieval Augmentation Reduces Hallucination in Conversation
Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela and Jason Weston .............. 3784

Towards Automatic Bias Detection in Knowledge Graphs
Daphna Keidar, Mian Zhong, Ce Zhang, Yash Raj Shrestha and Bibek Paudel .......... 3804

Searching for More Efficient Dynamic Programs
Tim Vieira, Ryan Cotterell and Jason Eisner ......................................................... 3812

Revisiting Robust Neural Machine Translation: A Transformer Case Study
Peyman Passban, Puneeth Saladi and Qun Liu ...................................................... 3831

Can NLI Models Verify QA Systems’ Predictions?
Jifan Chen, Eunsol Choi and Greg Durrett .......................................................... 3841

Parameter-Efficient Domain Knowledge Integration from Multiple Sources for Biomedical Pre-trained Language Models
Qiuhao Lu, Dejing Dou and Thien Huu Nguyen ..................................................... 3855

Uncovering Implicit Gender Bias in Narratives through Commonsense Inference
Tenghao Huang, Faeze Brahman, Vered Shwartz and Snigdha Chaturvedi ................ 3866

Contrastive Document Representation Learning with Graph Attention Networks
Peng Xu, Xinchi Chen, Xiaofei Ma, Zhiheng Huang and Bing Xiang ......................... 3874

Convex Aggregation for Opinion Summarization
Hayate Iso, Xiaolan Wang, Yoshihiko Suhara, Stefanos Angelidis and Wang-Chiew Tan ..... 3885

Using Optimal Transport as Alignment Objective for fine-tuning Multilingual Contextualized Embeddings
Sawsan Alqahtani, Garima Lalwani, Yi Zhang, Salvatore Romeo and Saab Mansour .... 3904

Uncertainty-Aware Machine Translation Evaluation
Taisiya Glushkova, Chrysoula Zerva, Ricardo Rei and André F. T. Martins ................ 3920

Neural Unification for Logic Reasoning over Natural Language
Gabriele Picco, Thanh Lam Hoang, Marco Luca Sbodio and Vanessa Lopez ............... 3939

From None to Severe: Predicting Severity in Movie Scripts
Yigeng Zhang, Mahsa Shafaei, Fabio Gonzalez and Thamar Solorio .......................... 3951

Benchmarking Meta-embeddings: What Works and What Does Not
Iker Garcia-Ferrero, Rodrigo Agerri and German Rigau ........................................ 3957

A Plug-and-Play Method for Controlled Text Generation
Damian Pascual, Beni Egressy, Clara Meister, Ryan Cotterell and Roger Wattenhofer .... 3973

A Corpus-based Syntactic Analysis of Two-termed Unlike Coordination
Julie Kallini and Christiane Fellbaum ................................................................. 3998
Weakly Supervised Contrastive Learning for Chest X-Ray Report Generation
An Yan, Zexue He, Xing Lu, Jiang Du, Eric Chang, Amilcare Gentili, Julian McAuley and Chun-Nan Hsu

NUANCED: Natural Utterance Annotation for Nuanced Conversation with Estimated Distributions
Zhiyu Chen, Honglei Liu, Hu Xu, Seungwan Moon, Hao Zhou and Bing Liu

Table-based Fact Verification With Salience-aware Learning
Fei Wang, Kexuan Sun, Jay Pujara, Pedro Szekely and Muhao Chen

Detecting Frames in News Headlines and Lead Images in U.S. Gun Violence Coverage
Isidora Tourni, Lei Guo, Taufiq Husada Daryanto, Fabian Zhafransyah, Edward Edberg Halim, Mona Jalal, Boqi Chen, Sha Lai, Hengchang Hu, Margrit Betke, Prakash Ishwar and Derry Tanti Wijaya

Multi-task Learning to Enable Location Mention Identification in the Early Hours of a Crisis Event
Sarthak Khanal and Doina Caragea

Graph-Based Decoding for Task Oriented Semantic Parsing
Jeremy Cole, Nanjiang Jiang, Panupong Pasupat, Luheng He and Peter Shaw

Expected Validation Performance and Estimation of a Random Variable’s Maximum
Jesse Dodge, Suchin Gururangan, Dallas Card, Roy Schwartz and Noah A. Smith

How May I Help You? Using Neural Text Simplification to Improve Downstream NLP Tasks
Hoang Van, Zheng Tang and Mihai Surdeanu

Subformer: Exploring Weight Sharing for Parameter Efficiency in Generative Transformers
Machel Reid, Edison Marrese-Taylor and Yutaka Matsuo

Leveraging Information Bottleneck for Scientific Document Summarization
Jiaxin Ju, Ming Liu, Huan Yee Koh, Yuan Jin, Lan Du and Shirui Pan

Reconsidering the Past: Optimizing Hidden States in Language Models
Davis Yoshida and Kevin Gimpel

Attend, Memorize and Generate: Towards Faithful Table-to-Text Generation in Few Shots
Wenting Zhao, Ye Liu, Yao Wan and Philip Yu

ARCH: Efficient Adversarial Regularized Training with Caching
Simiao Zuo, Chen Liang, Haoming Jiang, Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen and Tuo Zhao

Probing Commonsense Explanation in Dialogue Response Generation
Pei Zhou, Pegah Jandaghi, Hyundong Cho, Bill Yuchen Lin, Jay Pujara and Xiang Ren

NOAHQA: Numerical Reasoning with Interpretable Graph Question Answering Dataset
Qiyuan Zhang, Lei Wang, Sicheng Yu, Shuohang Wang, Yang Wang, Jing Jiang and Ee-Peng Lim

Textual Time Travel: A Temporally Informed Approach to Theory of Mind
Akshatha Arodi and Jackie Chi Kit Cheung

Detect and Perturb: Neutral Rewriting of Biased and Sensitive Text via Gradient-based Decoding
Zexue He, Bodhisattwa Prasad Majumder and Julian McAuley
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Character-based PCFG Induction for Modeling the Syntactic Acquisition of Morphologically Rich Languages</td>
<td>Lifeng Jin, Byung-Doh Oh and William Schuler</td>
<td>4367</td>
</tr>
<tr>
<td>Block-wise Word Embedding Compression Revisited: Better Weighting and Structuring</td>
<td>Jong-Ryul Lee, Yong-Ju Lee and Yong-Hyuk Moon</td>
<td>4379</td>
</tr>
<tr>
<td>Switch Point biased Self-Training: Re-purposing Pretrained Models for Code-Switching</td>
<td>Parul Chopra, Sai Krishna Rallabandi, Alan W Black and Khyathi Raghavi Chandu</td>
<td>4389</td>
</tr>
<tr>
<td>Influence Tuning: Demoting Spurious Correlations via Instance Attribution and Instance-Driven Updates</td>
<td>Xiaochuang Han and Yulia Tsvetkov</td>
<td>4398</td>
</tr>
<tr>
<td>Learning Task Sampling Policy for Multi-task Learning</td>
<td>Dhanasekar Sundararaman, Henry Tsai, Kuang-Huei Lee, Iulia Turc and Lawrence Carin</td>
<td>4410</td>
</tr>
<tr>
<td>Competing Independent Modules for Knowledge Integration and Optimization</td>
<td>Parsa Bagherzadeh and Sabine Bergler</td>
<td>4416</td>
</tr>
<tr>
<td>Improving Text Auto-Completion with Next Phrase Prediction</td>
<td>Dong-Ho Lee, Zhiqiang Hu and Roy Ka-Wei Lee</td>
<td>4434</td>
</tr>
<tr>
<td>MOMENTA: A Multimodal Framework for Detecting Harmful Memes and Their Targets</td>
<td>Shraman Pramanick, Shivam Sharma, Dimitar Dimitrov, Md. Shad Akhtar, Preslav Nakov and Tanmoy Chakraborty</td>
<td>4439</td>
</tr>
<tr>
<td>NICE: Neural Image Commenting with Empathy</td>
<td>Kezhen Chen, Qiuyuan Huang, Daniel McDuff, Xiang Gao, Hamid Palangi, Jianfeng Wang, Kenneth Forbus and Jianfeng Gao</td>
<td>4456</td>
</tr>
<tr>
<td>HACovGNN: Hierarchical Attention Based Convolutional Graph Neural Network for Code Documentation Generation in Jupyter Notebooks</td>
<td>Xuye Liu, Dakuo Wang, April Wang, Yufang Hou and Lingfei Wu</td>
<td>4473</td>
</tr>
<tr>
<td>A multilabel approach to morphosyntactic probing</td>
<td>Naomi Shapiro, Amandalynne Paullada and Shane Steinert-Threlkeld</td>
<td>4486</td>
</tr>
<tr>
<td>Co-Teaching Student-Model through Submission Results of Shared Task</td>
<td>Kouta Nakayama, Shuhei Kurita, Akio Kobayashi, Yukino Baba and Satoshi Sekine</td>
<td>4525</td>
</tr>
<tr>
<td>KLMo: Knowledge Graph Enhanced Pretrained Language Model with Fine-Grained Relationships</td>
<td>Lei He, Suncong Zheng, Tao Yang and Feng Zhang</td>
<td>4536</td>
</tr>
<tr>
<td>Do We Know What We Don’t Know? Studying Unanswerable Questions beyond SQuAD 2.0</td>
<td>Elior Sulem, Jamaal Hay and Dan Roth</td>
<td>4543</td>
</tr>
<tr>
<td>Glyph Enhanced Chinese Character Pre-Training for Lexical Sememe Prediction</td>
<td>Boer Lyu, Lu Chen and Kai Yu</td>
<td>4549</td>
</tr>
<tr>
<td>Active Learning for Rumor Identification on Social Media</td>
<td>Parsa Farinneya, Mohammad Mahdi Abdollah Pour, Sardar Hamidian and Mona Diab</td>
<td>4556</td>
</tr>
</tbody>
</table>
Don’t Discard All the Biased Instances: Investigating a Core Assumption in Dataset Bias Mitigation Techniques
Hossein Amirkhani and Mohammad Taher Pilehvar ........................................ 4720

Stacked AMR Parsing with Silver Data
Qingrong Xia, Zhenghua Li, Rui Wang and Min Zhang ........................................ 4729

Speculative Sampling in Variational Autoencoders for Dialogue Response Generation
Shoetsu Sato, Naoki Yoshinaga, Masashi Toyoda and Masaru Kitsuregawa .................. 4739

Perceived and Intended Sarcasm Detection with Graph Attention Networks
Joan Plepi and Lucie Flek ............................................................... 4746

Contrastive Representation Learning for Exemplar-Guided Paraphrase Generation
Haoran Yang, Wai Lam and Piji Li ....................................................... 4754

MAD-G: Multilingual Adapter Generation for Efficient Cross-Lingual Transfer
Alan Ansell, Edoardo Maria Ponti, Jonas Pfeiffer, Sebastian Ruder, Goran Glavaš, Ivan Vulić and Anna Korhonen ................................................................. 4762

Sustainable Modular Debiasing of Language Models
Anne Lauscher, Tobias Lueken and Goran Glavaš ............................................... 4782

A Divide-And-Conquer Approach for Multi-label Multi-hop Relation Detection in Knowledge Base Question Answering
Deyu Zhou, Yanzheng Xiang, Linhai Zhang, Chenchen Ye, Qian-Wen Zhang and Yunbo Cao 4798

Counterfactual Adversarial Learning with Representation Interpolation
Wei Wang, Boxin Wang, Ning Shi, Jinfeng Li, Bingyu Zhu, Xiangyu Liu and Rong Zhang ... 4809

‘Just What do You Think You’re Doing, Dave?’ A Checklist for Responsible Data Use in NLP
Anna Rogers, Timothy Baldwin and kobi leins ............................................... 4821

Counter-Contrastive Learning for Language GANs
Yekun Chai, Haidong Zhang, Qiuye Yin and Junge Zhang ................................... 4834

Incorporating Circumstances into Narrative Event Prediction
Shichao Wang, Xiangrui Cai, HongBin Wang and Xiaojie Yuan ............................... 4840

MultiFix: Learning to Repair Multiple Errors by Optimal Alignment Learning
HyeonTae Seo, Yo-Sub Han and Sang-Ki Ko ................................................. 4850

HOTTER: Hierarchical Optimal Topic Transport with Explanatory Context Representations
Sabine Wehnert, Christian Scheel, Simona Szakács-Behling, Maret Nieländer, Patrick Mielke and Ernesto William De Luca ............................................. 4856

Grammatical Error Correction with Contrastive Learning in Low Error Density Domains
Hannan Cao, Wenmian Yang and Hwee Tou Ng ............................................. 4867

Improving Unsupervised Commonsense Reasoning Using Knowledge-Enabled Natural Language Inference
Canming Huang, Weinan He and Yongmei Liu .............................................. 4875

Does Putting a Linguist in the Loop Improve NLU Data Collection?
Alicia Parrish, William Huang, Omar Agha, Soo-Hwan Lee, Nikita Nangia, Alexia Warstadt, Kar- manya Aggarwal, Emily Allaway, Tal Linzen and Samuel R. Bowman ...................... 4886

xxix
Tiered Reasoning for Intuitive Physics: Toward Verifiable Commonsense Language Understanding
Shane Storks, Qiaozi Gao, Yichi Zhang and Joyce Chai ........................................... 4902

Making Heads and Tails of Models with Marginal Calibration for Sparse Tagsets
Michael Kranzlein, Nelson F. Liu and Nathan Schneider........................................... 4919

GeDi: Generative Discriminator Guided Sequence Generation