57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)

Florence, Italy
28 July – 2 August 2019

Volume 1 of 10

ISBN: 978-1-5108-9096-1

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2019) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2019)

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

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

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

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

Table of Contents

One Time of Interaction May Not Be Enough: Go Deep with an Interaction-over-Interaction Network for Response Selection in Dialogues
Chongyang Tao, Wei Wu, Can Xu, Wenpeng Hu, Dongyan Zhao and Rui Yan 1

Incremental Transformer with Deliberation Decoder for Document Grounded Conversations
Zekang Li, Cheng Niu, Fandong Meng, Yang Feng, Qian Li and Jie Zhou 12

Improving Multi-turn Dialogue Modelling with Utterance ReWriter
Hui Su, Xiaoyu Shen, Rongzhi Zhang, Fei Sun, Pengwei Hu, Cheng Niu and Jie Zhou 22

Do Neural Dialog Systems Use the Conversation History Effectively? An Empirical Study
Chinnadhurai Sankar, Sandeep Subramanian, Chris Pal, Sarath Chandar and Yoshua Bengio . . . 32

Boosting Dialog Response Generation
Wenchao Du and Alan W Black 38

Constructing Interpretive Spatio-Temporal Features for Multi-Turn Responses Selection
Junyu Lu, Chenbin Zhang, Zeying Xie, Guang Ling, Tom Chao Zhou and Zenglin Xu 44

Semantic Parsing with Dual Learning
Ruisheng Cao, Su Zhu, Chen Liu, Jieyu Li and Kai Yu 51

Semantic Expressive Capacity with Bounded Memory
Antoine Venant and Alexander Koller 65

AMR Parsing as Sequence-to-Graph Transduction
Sheng Zhang, Xutai Ma, Kevin Duh and Benjamin Van Durme 80

Generating Logical Forms from Graph Representations of Text and Entities
Peter Shaw, Philip Massey, Angelica Chen, Francesco Piccinno and Yasemin Altun 95

Learning Compressed Sentence Representations for On-Device Text Processing
Dinghan Shen, Pengyu Cheng, Dhanasekar Sundararaman, Xinyuan Zhang, Qian Yang, Meng Tang, Asli Celikyilmaz and Lawrence Carin 107

The (Non-)Utility of Structural Features in BiLSTM-based Dependency Parsers
Agnieszka Falenska and Jonas Kuhn 117

Automatic Generation of High Quality CCGbanks for Parser Domain Adaptation
Masashi Yoshikawa, Hiroshi Noji, Koji Mineshima and Daisuke Bekki 129

A Joint Named-Entity Recognizer for Heterogeneous Tag-sets Using a Tag Hierarchy
Genady Beryozkin, Yoel Drori, Oren Gilon, Tzvika Hartman and Idan Szpektor 140

Massively Multilingual Transfer for NER
Afshin Rahimi, Yuan Li and Trevor Cohn 151

Reliability-aware Dynamic Feature Composition for Name Tagging
Ying Lin, Liyuan Liu, Heng Ji, Dong Yu and Jiawei Han 165

Unsupervised Pivot Translation for Distant Languages
Yichong Leng, Xu Tan, Tao Qin, Xiang-Yang Li and Tie-Yan Liu 175

Bilingual Lexicon Induction with Semi-supervision in Non-Isometric Embedding Spaces
Barun Patra, Joel Ruben Antony Moniz, Sarthak Garg, Matthew R. Gormley and Graham Neubig 184

An Effective Approach to Unsupervised Machine Translation
Mikel Artetxe, Gorka Labaka and Eneko Agirre 194

Effective Adversarial Regularization for Neural Machine Translation
Motoki Sato, Jun Suzuki and Shun Kiyono 204

Revisiting Low-Resource Neural Machine Translation: A Case Study
Rico Sennrich and Biao Zhang 211

Domain Adaptive Inference for Neural Machine Translation
Danielle Saunders, Felix Stahlberg, Adrià de Gispert and Bill Byrne 222

Neural Relation Extraction for Knowledge Base Enrichment
Bayu Distiawan Trisedya, Gerhard Weikum, Jianzhong Qi and Rui Zhang 229

Attention Guided Graph Convolutional Networks for Relation Extraction
Zhijiang Guo, Yan Zhang and Wei Lu 241

Spatial Aggregation Facilitates Discovery of Spatial Topics
Aniruddha Maiti and Slobodan Vucetic 252

Relation Embedding with Dihedral Group in Knowledge Graph
Canran Xu and Ruijiang Li 263

Sequence Tagging with Contextual and Non-Contextual Subword Representations: A Multilingual Evaluation
Benjamin Heinzerling and Michael Strube 273

Augmenting Neural Networks with First-order Logic
Tao Li and Vivek Srikumar 292

Self-Regulated Interactive Sequence-to-Sequence Learning
Julia Kreutzer and Stefan Riezler 303

You Only Need Attention to Traverse Trees
Mahtab Ahmed, Muhammad Rifayat Samee and Robert E. Mercer 316

Cross-Domain Generalization of Neural Constituency Parsers
Daniel Fried, Nikita Kitaev and Dan Klein 323

Adaptive Attention Span in Transformers
Sainbayar Sukhbaatar, Edouard Grave, Piotr Bojanowski and Armand Joulin 331

Neural News Recommendation with Long- and Short-term User Representations
Mingxiao An, Fangzhao Wu, Chuhan Wu, Kun Zhang, Zheng Liu and Xing Xie 336

Automatic Domain Adaptation Outperforms Manual Domain Adaptation for Predicting Financial Outcomes
Marina Sedinkina, Nikolas Breitkopf and Hinrich Schütze 346

Manipulating the Difficulty of C-Tests
Ji-Ung Lee, Erik Schwan and Christian M. Meyer 360

Towards Unsupervised Text Classification Leveraging Experts and Word Embeddings
Zied Haj-Yahia, Adrien Sieg and Léa A. Deleris 371

Neural Text Simplification of Clinical Letters with a Domain Specific Phrase Table
Matthew Shardlow and Raheel Nawaz 380

What You Say and How You Say It Matters: Predicting Stock Volatility Using Verbal and Vocal Cues
Yu Qin and Yi Yang 390

Detecting Concealed Information in Text and Speech
Shengli Hu 402

Evidence-based Trustworthiness
Yi Zhang, Zachary Ives and Dan Roth 413

Disentangled Representation Learning for Non-Parallel Text Style Transfer
Vineet John, Lili Mou, Hareesh Bahuleyan and Olga Vechtomova 424

Cross-Sentence Grammatical Error Correction
Shamil Chollampatt, Weiqi Wang and Hwee Tou Ng 435

This Email Could Save Your Life: Introducing the Task of Email Subject Line Generation
Rui Zhang and Joel Tetreault 446

Time-Out: Temporal Referencing for Robust Modeling of Lexical Semantic Change
Haim Dubossarsky, Simon Hengchen, Nina Tahmasebi and Dominik Schlechtweg 457

Adversarial Attention Modeling for Multi-dimensional Emotion Regression
Suyang Zhu, Shoushan Li and Guodong Zhou 471

Divide, Conquer and Combine: Hierarchical Feature Fusion Network with Local and Global Perspectives for Multimodal Affective Computing
Sijie Mai, Haifeng Hu and Songlong Xing 481

Modeling Financial Analysts' Decision Making via the Pragmatics and Semantics of Earnings Calls
Katherine Keith and Amanda Stent 493

An Interactive Multi-Task Learning Network for End-to-End Aspect-Based Sentiment Analysis
Ruidan He, Wee Sun Lee, Hwee Tou Ng and Daniel Dahlmeier 504

Decompositional Argument Mining: A General Purpose Approach for Argument Graph Construction
Debela Gemechu and Chris Reed 516

MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversations
Soujanya Poria, Devamanyu Hazarika, Navonil Majumder, Gautam Naik, Erik Cambria and Rada Mihalcea 527

Open-Domain Targeted Sentiment Analysis via Span-Based Extraction and Classification
Minghao Hu, Yuxing Peng, Zhen Huang, Dongsheng Li and Yiwei Lv 537

Transfer Capsule Network for Aspect Level Sentiment Classification
Zhuang Chen and Tieyun Qian 547

Progressive Self-Supervised Attention Learning for Aspect-Level Sentiment Analysis
Jialong Tang, Ziyao Lu, Jinsong Su, Yubin Ge, Linfeng Song, Le Sun and Jiebo Luo 557

Classification and Clustering of Arguments with Contextualized Word Embeddings
Nils Reimers, Benjamin Schiller, Tilman Beck, Johannes Daxenberger, Christian Stab and Iryna Gurevych 567

Sentiment Tagging with Partial Labels using Modular Architectures
Xiao Zhang and Dan Goldwasser 579

DOER: Dual Cross-Shared RNN for Aspect Term-Polarity Co-Extraction
Huaishao Luo, Tianrui Li, Bing Liu and Junbo Zhang 591

A Corpus for Modeling User and Language Effects in Argumentation on Online Debating
Esin Durmus and Claire Cardie 602

Topic Tensor Network for Implicit Discourse Relation Recognition in Chinese
Sheng Xu, Peifeng Li, Fang Kong, Qiaoming Zhu and Guodong Zhou 608

Learning from Omission
Bill McDowell and Noah Goodman 619

Multi-Task Learning for Coherence Modeling
Youmna Farag and Helen Yannakoudakis 629

Data Programming for Learning Discourse Structure
Sonia Badene, Kate Thompson, Jean-Pierre Lorré and Nicholas Asher 640

Evaluating Discourse in Structured Text Representations
Elisa Ferracane, Greg Durrett, Junyi Jessy Li and Katrin Erk 646

Know What You Don't Know: Modeling a Pragmatic Speaker that Refers to Objects of Unknown Categories
Sina Zarrieß and David Schlangen 654

End-to-end Deep Reinforcement Learning Based Coreference Resolution
Hongliang Fei, Xu Li, Dingcheng Li and Ping Li 660

Implicit Discourse Relation Identification for Open-domain Dialogues
Mingyu Derek Ma, Kevin Bowden, Jiaqi Wu, Wen Cui and Marilyn Walker 666

Coreference Resolution with Entity Equalization
Ben Kantor and Amir Globerson 673

A Cross-Domain Transferable Neural Coherence Model
Peng Xu, Hamidreza Saghir, Jin Sung Kang, Teng Long, Avishek Joey Bose, Yanshuai Cao and Jackie Chi Kit Cheung 678

MOROCO: The Moldavian and Romanian Dialectal Corpus
Andrei Butnaru and Radu Tudor Ionescu 688

Just "OneSeC" for Producing Multilingual Sense-Annotated Data
Bianca Scarlini, Tommaso Pasini and Roberto Navigli 699

How to (Properly) Evaluate Cross-Lingual Word Embeddings: On Strong Baselines, Comparative Analyses, and Some Misconceptions
Goran Glavaš, Robert Litschko, Sebastian Ruder and Ivan Vulić 710

SP-10K: A Large-scale Evaluation Set for Selectional Preference Acquisition
Hongming Zhang, Hantian Ding and Yangqiu Song 722

A Wind of Change: Detecting and Evaluating Lexical Semantic Change across Times and Domains
Dominik Schlechtweg, Anna Hätty, Marco Del Tredici and Sabine Schulte im Walde 732

Errudite: Scalable, Reproducible, and Testable Error Analysis
Tongshuang Wu, Marco Tulio Ribeiro, Jeffrey Heer and Daniel Weld 747

DocRED: A Large-Scale Document-Level Relation Extraction Dataset
Yuan Yao, Deming Ye, Peng Li, Xu Han, Yankai Lin, Zhenghao Liu, Zhiyuan Liu, Lixin Huang, Jie Zhou and Maosong Sun 764

ChID: A Large-scale Chinese IDiom Dataset for Cloze Test
Chujie Zheng, Minlie Huang and Aixin Sun 778

Automatic Evaluation of Local Topic Quality
Jeffrey Lund, Piper Armstrong, Wilson Fearn, Stephen Cowley, Courtni Byun, Jordan Boyd-Graber and Kevin Seppi 788

Crowdsourcing and Aggregating Nested Markable Annotations
Chris Madge, Juntao Yu, Jon Chamberlain, Udo Kruschwitz, Silviu Paun and Massimo Poesio 797

Transferable Multi-Domain State Generator for Task-Oriented Dialogue Systems
Chien-Sheng Wu, Andrea Madotto, Ehsan Hosseini-Asl, Caiming Xiong, Richard Socher and Pascale Fung 808

Multi-Task Networks with Universe, Group, and Task Feature Learning
Shiva Pentyala, Mengwen Liu and Markus Dreyer 820

Constrained Decoding for Neural NLG from Compositional Representations in Task-Oriented Dialogue
Anusha Balakrishnan, Jinfeng Rao, Kartikeya Upasani, Michael White and Rajen Subba 831

OpenDialKG: Explainable Conversational Reasoning with Attention-based Walks over Knowledge Graphs
Seungwhan Moon, Pararth Shah, Anuj Kumar and Rajen Subba 845

Coupling Retrieval and Meta-Learning for Context-Dependent Semantic Parsing
Daya Guo, Duyu Tang, Nan Duan, Ming Zhou and Jian Yin 855

Knowledge-aware Pronoun Coreference Resolution
Hongming Zhang, Yan Song, Yangqiu Song and Dong Yu 867

Don't Take the Premise for Granted: Mitigating Artifacts in Natural Language Inference
Yonatan Belinkov, Adam Poliak, Stuart Shieber, Benjamin Van Durme and Alexander Rush . . . 877

GEAR: Graph-based Evidence Aggregating and Reasoning for Fact Verification
Jie Zhou, Xu Han, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li and Maosong Sun 892

SherLIiC: A Typed Event-Focused Lexical Inference Benchmark for Evaluating Natural Language Inference
Martin Schmitt and Hinrich Schütze 902

Extracting Symptoms and their Status from Clinical Conversations
Nan Du, Kai Chen, Anjuli Kannan, Linh Tran, Yuhui Chen and Izhak Shafran 915

What Makes a Good Counselor? Learning to Distinguish between High-quality and Low-quality Counseling Conversations
Verónica Pérez-Rosas, Xinyi Wu, Kenneth Resnicow and Rada Mihalcea 926

Finding Your Voice: The Linguistic Development of Mental Health Counselors
Justine Zhang, Robert Filbin, Christine Morrison, Jaclyn Weiser and Cristian Danescu-Niculescu-Mizil 936

Towards Automating Healthcare Question Answering in a Noisy Multilingual Low-Resource Setting
Jeanne E. Daniel, Willie Brink, Ryan Eloff and Charles Copley 948

Joint Entity Extraction and Assertion Detection for Clinical Text
Parminder Bhatia, Busra Celikkaya and Mohammed Khalilia 954

HEAD-QA: A Healthcare Dataset for Complex Reasoning
David Vilares and Carlos Gómez-Rodríguez 960

Are You Convinced? Choosing the More Convincing Evidence with a Siamese Network
Martin Gleize, Eyal Shnarch, Leshem Choshen, Lena Dankin, Guy Moshkowich, Ranit Aharonov and Noam Slonim 967

From Surrogacy to Adoption; From Bitcoin to Cryptocurrency: Debate Topic Expansion
Roy Bar-Haim, Dalia Krieger, Orith Toledo-Ronen, Lilach Edelstein, Yonatan Bilu, Alon Halfon, Yoav Katz, Amir Menczel, Ranit Aharonov and Noam Slonim 977

Multimodal and Multi-view Models for Emotion Recognition
Gustavo Aguilar, Viktor Rozgic, Weiran Wang and Chao Wang 991

Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts
Rui Xia and Zixiang Ding 1003

Argument Invention from First Principles
Yonatan Bilu, Ariel Gera, Daniel Hershcovich, Benjamin Sznajder, Dan Lahav, Guy Moshkowich, Anael Malet, Assaf Gavron and Noam Slonim 1013

Improving the Similarity Measure of Determinantal Point Processes for Extractive Multi-Document Summarization
Sangwoo Cho, Logan Lebanoff, Hassan Foroosh and Fei Liu 1027

Global Optimization under Length Constraint for Neural Text Summarization
Takuya Makino, Tomoya Iwakura, Hiroya Takamura and Manabu Okumura 1039

Searching for Effective Neural Extractive Summarization: What Works and What's Next
Ming Zhong, Pengfei Liu, Danqing Wang, Xipeng Qiu and Xuanjing Huang 1049

A Simple Theoretical Model of Importance for Summarization
Maxime Peyrard 1059

Multi-News: A Large-Scale Multi-Document Summarization Dataset and Abstractive Hierarchical Model
Alexander Fabbri, Irene Li, Tianwei She, Suyi Li and Dragomir Radev 1074

Generating Natural Language Adversarial Examples through Probability Weighted Word Saliency
Shuhuai Ren, Yihe Deng, Kun He and Wanxiang Che 1085

Heuristic Authorship Obfuscation
Janek Bevendorff, Martin Potthast, Matthias Hagen and Benno Stein 1098

Text Categorization by Learning Predominant Sense of Words as Auxiliary Task
Kazuya Shimura, Jiyi Li and Fumiyo Fukumoto 1109

DeepSentiPeer: Harnessing Sentiment in Review Texts to Recommend Peer Review Decisions
Tirthankar Ghosal, Rajeev Verma, Asif Ekbal and Pushpak Bhattacharyya 1120

Gated Embeddings in End-to-End Speech Recognition for Conversational-Context Fusion
Suyoun Kim, Siddharth Dalmia and Florian Metze 1131

Figurative Usage Detection of Symptom Words to Improve Personal Health Mention Detection
Adith Iyer, Aditya Joshi, Sarvnaz Karimi, Ross Sparks and Cecile Paris 1142

Complex Word Identification as a Sequence Labelling Task
Sian Gooding and Ekaterina Kochmar 1148

Neural News Recommendation with Topic-Aware News Representation
Chuhan Wu, Fangzhao Wu, Mingxiao An, Yongfeng Huang and Xing Xie 1154

Poetry to Prose Conversion in Sanskrit as a Linearisation Task: A Case for Low-Resource Languages
Amrith Krishna, Vishnu Sharma, Bishal Santra, Aishik Chakraborty, Pavankumar Satuluri and Pawan Goyal 1160

Learning Emphasis Selection for Written Text in Visual Media from Crowd-Sourced Label Distributions
Amirreza Shirani, Franck Dernoncourt, Paul Asente, Nedim Lipka, Seokhwan Kim, Jose Echevarria and Thamar Solorio 1167

Rumor Detection by Exploiting User Credibility Information, Attention and Multi-task Learning
Quanzhi Li, Qiong Zhang and Luo Si 1173

Context-specific Language Modeling for Human Trafficking Detection from Online Advertisements
Saeideh Shahrokh Esfahani, Michael J. Cafarella, Maziyar Baran Pouyan, Gregory DeAngelo, Elena Eneva and Andy E. Fano 1180

Self-Attentional Models for Lattice Inputs
Matthias Sperber, Graham Neubig, Ngoc-Quan Pham and Alex Waibel 1185

When a Good Translation is Wrong in Context: Context-Aware Machine Translation Improves on Deixis, Ellipsis, and Lexical Cohesion
Elena Voita, Rico Sennrich and Ivan Titov 1198

A Compact and Language-Sensitive Multilingual Translation Method
Yining Wang, Long Zhou, Jiajun Zhang, Feifei Zhai, Jingfang Xu and Chengqing Zong 1213

Unsupervised Parallel Sentence Extraction with Parallel Segment Detection Helps Machine Translation
Viktor Hangya and Alexander Fraser 1224

Unsupervised Bilingual Word Embedding Agreement for Unsupervised Neural Machine Translation
Haipeng Sun, Rui Wang, Kehai Chen, Masao Utiyama, Eiichiro Sumita and Tiejun Zhao 1235

Effective Cross-lingual Transfer of Neural Machine Translation Models without Shared Vocabularies
Yunsu Kim, Yingbo Gao and Hermann Ney 1246

Improved Zero-shot Neural Machine Translation via Ignoring Spurious Correlations
Jiatao Gu, Yong Wang, Kyunghyun Cho and Victor O.K. Li 1258

Syntactically Supervised Transformers for Faster Neural Machine Translation
Nader Akoury, Kalpesh Krishna and Mohit Iyyer 1269

Dynamically Composing Domain-Data Selection with Clean-Data Selection by "Co-Curricular Learning" for Neural Machine Translation
Wei Wang, Isaac Caswell and Ciprian Chelba 1282

On the Word Alignment from Neural Machine Translation
Xintong Li, Guanlin Li, Lemao Liu, Max Meng and Shuming Shi 1293

Imitation Learning for Non-Autoregressive Neural Machine Translation
Bingzhen Wei, Mingxuan Wang, Hao Zhou, Junyang Lin and Xu Sun 1304

Monotonic Infinite Lookback Attention for Simultaneous Machine Translation
Naveen Arivazhagan, Colin Cherry, Wolfgang Macherey, Chung-Cheng Chiu, Semih Yavuz, Ruoming Pang, Wei Li and Colin Raffel 1313

Global Textual Relation Embedding for Relational Understanding
Zhiyu Chen, Hanwen Zha, Honglei Liu, Wenhu Chen, Xifeng Yan and Yu Su 1324

Graph Neural Networks with Generated Parameters for Relation Extraction
Hao Zhu, Yankai Lin, Zhiyuan Liu, Jie Fu, Tat-Seng Chua and Maosong Sun 1331

Entity-Relation Extraction as Multi-Turn Question Answering
Xiaoya Li, Fan Yin, Zijun Sun, Xiayu Li, Arianna Yuan, Duo Chai, Mingxin Zhou and Jiwei Li 1340

Exploiting Entity BIO Tag Embeddings and Multi-task Learning for Relation Extraction with Imbalanced Data
Wei Ye, Bo Li, Rui Xie, Zhonghao Sheng, Long Chen and Shikun Zhang 1351

Joint Type Inference on Entities and Relations via Graph Convolutional Networks
Changzhi Sun, Yeyun Gong, Yuanbin Wu, Ming Gong, Daxin Jiang, Man Lan, Shiliang Sun and Nan Duan 1361

Extracting Multiple-Relations in One-Pass with Pre-Trained Transformers
Haoyu Wang, Ming Tan, Mo Yu, Shiyu Chang, Dakuo Wang, Kun Xu, Xiaoxiao Guo and Saloni Potdar 1371

Unsupervised Information Extraction: Regularizing Discriminative Approaches with Relation Distribution Losses
Étienne Simon, Vincent Guigue and Benjamin Piwowarski 1378

Fine-tuning Pre-Trained Transformer Language Models to Distantly Supervised Relation Extraction
Christoph Alt, Marc Hübner and Leonhard Hennig 1388

ARNOR: Attention Regularization based Noise Reduction for Distant Supervision Relation Classification
Wei Jia, Dai Dai, Xinyan Xiao and Hua Wu 1399

GraphRel: Modeling Text as Relational Graphs for Joint Entity and Relation Extraction
Tsu-Jui Fu, Peng-Hsuan Li and Wei-Yun Ma 1409

DIAG-NRE: A Neural Pattern Diagnosis Framework for Distantly Supervised Neural Relation Extraction
Shun Zheng, Xu Han, Yankai Lin, Peilin Yu, Lu Chen, Ling Huang, Zhiyuan Liu and Wei Xu 1419

Multi-grained Named Entity Recognition
Congying Xia, Chenwei Zhang, Tao Yang, Yaliang Li, Nan Du, Xian Wu, Wei Fan, Fenglong Ma and Philip Yu 1430

ERNIE: Enhanced Language Representation with Informative Entities
Zhengyan Zhang, Xu Han, Zhiyuan Liu, Xin Jiang, Maosong Sun and Qun Liu 1441

Multi-Channel Graph Neural Network for Entity Alignment
Yixin Cao, Zhiyuan Liu, Chengjiang Li, Zhiyuan Liu, Juanzi Li and Tat-Seng Chua 1452

A Neural Multi-digraph Model for Chinese NER with Gazetteers
Ruixue Ding, Pengjun Xie, Xiaoyan Zhang, Wei Lu, Linlin Li and Luo Si 1462

Improved Language Modeling by Decoding the Past
Siddhartha Brahma 1468

Training Hybrid Language Models by Marginalizing over Segmentations
Edouard Grave, Sainbayar Sukhbaatar, Piotr Bojanowski and Armand Joulin 1477

Improving Neural Language Models by Segmenting, Attending, and Predicting the Future
Hongyin Luo, Lan Jiang, Yonatan Belinkov and James Glass 1483

Lightweight and Efficient Neural Natural Language Processing with Quaternion Networks
Yi Tay, Aston Zhang, Anh Tuan Luu, Jinfeng Rao, Shuai Zhang, Shuohang Wang, Jie Fu and Siu Cheung Hui 1494

Sparse Sequence-to-Sequence Models
Ben Peters, Vlad Niculae and André F. T. Martins 1504

On the Robustness of Self-Attentive Models
Yu-Lun Hsieh, Minhao Cheng, Da-Cheng Juan, Wei Wei, Wen-Lian Hsu and Cho-Jui Hsieh . 1520

Exact Hard Monotonic Attention for Character-Level Transduction
Shijie Wu and Ryan Cotterell 1530

A Lightweight Recurrent Network for Sequence Modeling
Biao Zhang and Rico Sennrich 1538

Towards Scalable and Reliable Capsule Networks for Challenging NLP Applications
Wei Zhao, Haiyun Peng, Steffen Eger, Erik Cambria and Min Yang 1549

Soft Representation Learning for Sparse Transfer
Haeju Park, Jinyoung Yeo, Gengyu Wang and Seung-won Hwang 1560

Learning Representations from Imperfect Time Series Data via Tensor Rank Regularization
Paul Pu Liang, Zhun Liu, Yao-Hung Hubert Tsai, Qibin Zhao, Ruslan Salakhutdinov and Louis-Philippe Morency 1569

Towards Lossless Encoding of Sentences
Gabriele Prato, Mathieu Duchesneau, Sarath Chandar and Alain Tapp 1577

Open Vocabulary Learning for Neural Chinese Pinyin IME
Zhuosheng Zhang, Yafang Huang and Hai Zhao 1584

Using LSTMs to Assess the Obligatoriness of Phonological Distinctive Features for Phonotactic Learning
Nicole Mirea and Klinton Bicknell 1595

Better Character Language Modeling through Morphology
Terra Blevins and Luke Zettlemoyer 1606

Historical Text Normalization with Delayed Rewards
Simon Flachs, Marcel Bollmann and Anders Søgaard 1614

Stochastic Tokenization with a Language Model for Neural Text Classification
Tatsuya Hiraoka, Hiroyuki Shindo and Yuji Matsumoto 1620

Mitigating Gender Bias in Natural Language Processing: Literature Review
Tony Sun, Andrew Gaut, Shirlyn Tang, Yuxin Huang, Mai ElSherief, Jieyu Zhao, Diba Mirza, Elizabeth Belding, Kai-Wei Chang and William Yang Wang 1630

Gender-preserving Debiasing for Pre-trained Word Embeddings
Masahiro Kaneko and Danushka Bollegala 1641

Counterfactual Data Augmentation for Mitigating Gender Stereotypes in Languages with Rich Morphology
Ran Zmigrod, Sebastian J. Mielke, Hanna Wallach and Ryan Cotterell 1651

A Transparent Framework for Evaluating Unintended Demographic Bias in Word Embeddings
Chris Sweeney and Maryam Najafian 1662

The Risk of Racial Bias in Hate Speech Detection
Maarten Sap, Dallas Card, Saadia Gabriel, Yejin Choi and Noah A. Smith 1668

Evaluating Gender Bias in Machine Translation
Gabriel Stanovsky, Noah A. Smith and Luke Zettlemoyer 1679

LSTMEmbed: Learning Word and Sense Representations from a Large Semantically Annotated Corpus with Long Short-Term Memories
Ignacio Iacobacci and Roberto Navigli 1685

Understanding Undesirable Word Embedding Associations
Kawin Ethayarajh, David Duvenaud and Graeme Hirst 1696

Unsupervised Discovery of Gendered Language through Latent-Variable Modeling
Alexander Miserlis Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, Isabelle Augenstein and Ryan Cotterell 1706

Topic Sensitive Attention on Generic Corpora Corrects Sense Bias in Pretrained Embeddings
Vihari Piratla, Sunita Sarawagi and Soumen Chakrabarti 1717

SphereRE: Distinguishing Lexical Relations with Hyperspherical Relation Embeddings
Chengyu Wang, Xiaofeng He and Aoying Zhou 1727

Multilingual Factor Analysis
Francisco Vargas, Kamen Brestnichki, Alex Papadopoulos Korfiatis and Nils Hammerla 1738

Meaning to Form: Measuring Systematicity as Information
Tiago Pimentel, Arya D. McCarthy, Damian Blasi, Brian Roark and Ryan Cotterell 1751

Learning Morphosyntactic Analyzers from the Bible via Iterative Annotation Projection across 26 Languages
Garrett Nicolai and David Yarowsky 1765

Adversarial Multitask Learning for Joint Multi-Feature and Multi-Dialect Morphological Modeling
Nasser Zalmout and Nizar Habash 1775

Neural Machine Translation with Reordering Embeddings
Kehai Chen, Rui Wang, Masao Utiyama and Eiichiro Sumita 1787

Neural Fuzzy Repair: Integrating Fuzzy Matches into Neural Machine Translation
Bram Bulte and Arda Tezcan 1800

Learning Deep Transformer Models for Machine Translation
Qiang Wang, Bei Li, Tong Xiao, Jingbo Zhu, Changliang Li, Derek F. Wong and Lidia S. Chao 1810

Generating Diverse Translations with Sentence Codes
Raphael Shu, Hideki Nakayama and Kyunghyun Cho 1823

Self-Supervised Neural Machine Translation
Dana Ruiter, Cristina España-Bonet and Josef van Genabith 1828

Exploring Phoneme-Level Speech Representations for End-to-End Speech Translation
Elizabeth Salesky, Matthias Sperber and Alan W Black 1835

Visually Grounded Neural Syntax Acquisition
Haoyue Shi, Jiayuan Mao, Kevin Gimpel and Karen Livescu 1842

Stay on the Path: Instruction Fidelity in Vision-and-Language Navigation
Vihan Jain, Gabriel Magalhaes, Alexander Ku, Ashish Vaswani, Eugene Ie and Jason Baldridge 1862

Expressing Visual Relationships via Language
Hao Tan, Franck Dernoncourt, Zhe Lin, Trung Bui and Mohit Bansal 1873

Weakly-Supervised Spatio-Temporally Grounding Natural Sentence in Video
Zhenfang Chen, Lin Ma, Wenhan Luo and Kwan-Yee Kenneth Wong 1884

The PhotoBook Dataset: Building Common Ground through Visually-Grounded Dialogue
Janosch Haber, Tim Baumgärtner, Ece Takmaz, Lieke Gelderloos, Elia Bruni and Raquel Fernández 1895

Continual and Multi-Task Architecture Search
Ramakanth Pasunuru and Mohit Bansal 1911

Semi-supervised Stochastic Multi-Domain Learning using Variational Inference
Yitong Li, Timothy Baldwin and Trevor Cohn 1923

Boosting Entity Linking Performance by Leveraging Unlabeled Documents
Phong Le and Ivan Titov 1935

Pre-Learning Environment Representations for Data-Efficient Neural Instruction Following
David Gaddy and Dan Klein 1946

Reinforced Training Data Selection for Domain Adaptation
Miaofeng Liu, Yan Song, Hongbin Zou and Tong Zhang 1957

Generating Long and Informative Reviews with Aspect-Aware Coarse-to-Fine Decoding
Junyi Li, Wayne Xin Zhao, Ji-Rong Wen and Yang Song 1969

PaperRobot: Incremental Draft Generation of Scientific Ideas
Qingyun Wang, Lifu Huang, Zhiying Jiang, Kevin Knight, Heng Ji, Mohit Bansal and Yi Luan 1980

Rhetorically Controlled Encoder-Decoder for Modern Chinese Poetry Generation
Zhiqiang Liu, Zuohui Fu, Jie Cao, Gerard de Melo, Yik-Cheung Tam, Cheng Niu and Jie Zhou 1992

Enhancing Topic-to-Essay Generation with External Commonsense Knowledge
Pengcheng Yang, Lei Li, Fuli Luo, Tianyu Liu and Xu Sun 2002

Towards Fine-grained Text Sentiment Transfer
Fuli Luo, Peng Li, Pengcheng Yang, Jie Zhou, Yutong Tan, Baobao Chang, Zhifang Sui and Xu Sun 2013

Data-to-text Generation with Entity Modeling
Ratish Puduppully, Li Dong and Mirella Lapata 2023

Ensuring Readability and Data-fidelity using Head-modifier Templates in Deep Type Description Generation
Jiangjie Chen, Ao Wang, Haiyun Jiang, Suo Feng, Chenguang Li and Yanghua Xiao 2036

Key Fact as Pivot: A Two-Stage Model for Low Resource Table-to-Text Generation
Shuming Ma, Pengcheng Yang, Tianyu Liu, Peng Li, Jie Zhou and Xu Sun 2047

Unsupervised Neural Text Simplification
Sai Surya, Abhijit Mishra, Anirban Laha, Parag Jain and Karthik Sankaranarayanan 2058

Syntax-Infused Variational Autoencoder for Text Generation
Xinyuan Zhang, Yi Yang, Siyang Yuan, Dinghan Shen and Lawrence Carin 2069

Towards Generating Long and Coherent Text with Multi-Level Latent Variable Models
Dinghan Shen, Asli Celikyilmaz, Yizhe Zhang, Liqun Chen, Xin Wang, Jianfeng Gao and Lawrence Carin 2079

Jointly Learning Semantic Parser and Natural Language Generator via Dual Information Maximization
Hai Ye, Wenjie Li and Lu Wang 2090

Learning to Select, Track, and Generate for Data-to-Text
Hayate Iso, Yui Uehara, Tatsuya Ishigaki, Hiroshi Noji, Eiji Aramaki, Ichiro Kobayashi, Yusuke Miyao, Naoaki Okazaki and Hiroya Takamura 2102

Reinforced Dynamic Reasoning for Conversational Question Generation
Boyuan Pan, Hao Li, Ziyu Yao, Deng Cai and Huan Sun 2114

TalkSumm: A Dataset and Scalable Annotation Method for Scientific Paper Summarization Based on Conference Talks
Guy Lev, Michal Shmueli-Scheuer, Jonathan Herzig, Achiya Jerbi and David Konopnicki . . . 2125

Improving Abstractive Document Summarization with Salient Information Modeling
Yongjian You, Weijia Jia, Tianyi Liu and Wenmian Yang 2132

Unsupervised Neural Single-Document Summarization of Reviews via Learning Latent Discourse Structure and its Ranking
Masaru Isonuma, Junichiro Mori and Ichiro Sakata 2142

BiSET: Bi-directional Selective Encoding with Template for Abstractive Summarization
Kai Wang, Xiaojun Quan and Rui Wang 2153

Neural Keyphrase Generation via Reinforcement Learning with Adaptive Rewards
Hou Pong Chan, Wang Chen, Lu Wang and Irwin King 2163

Scoring Sentence Singletons and Pairs for Abstractive Summarization
Logan Lebanoff, Kaiqiang Song, Franck Dernoncourt, Doo Soon Kim, Seokhwan Kim, Walter Chang and Fei Liu 2175

Keep Meeting Summaries on Topic: Abstractive Multi-Modal Meeting Summarization
Manling Li, Lingyu Zhang, Heng Ji and Richard J. Radke 2190

Adversarial Domain Adaptation Using Artificial Titles for Abstractive Title Generation
Francine Chen and Yan-Ying Chen 2197

BIGPATENT: A Large-Scale Dataset for Abstractive and Coherent Summarization
Eva Sharma, Chen Li and Lu Wang 2204

Ranking Generated Summaries by Correctness: An Interesting but Challenging Application for Natural Language Inference
Tobias Falke, Leonardo F. R. Ribeiro, Prasetya Ajie Utama, Ido Dagan and Iryna Gurevych . . 2214

Self-Supervised Learning for Contextualized Extractive Summarization
Hong Wang, Xin Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang and William Yang Wang 2221

On the Summarization of Consumer Health Questions
Asma Ben Abacha and Dina Demner-Fushman 2228

Unsupervised Rewriter for Multi-Sentence Compression
Yang Zhao, Xiaoyu Shen, Wei Bi and Akiko Aizawa 2235

Inferential Machine Comprehension: Answering Questions by Recursively Deducing the Evidence Chain from Text
Jianxing Yu, Zhengjun Zha and Jian Yin 2241

Token-level Dynamic Self-Attention Network for Multi-Passage Reading Comprehension
Yimeng Zhuang and Huadong Wang 2252

Explicit Utilization of General Knowledge in Machine Reading Comprehension
Chao Wang and Hui Jiang 2263

Multi-style Generative Reading Comprehension
Kyosuke Nishida, Itsumi Saito, Kosuke Nishida, Kazutoshi Shinoda, Atsushi Otsuka, Hisako Asano and Junji Tomita 2273

Retrieve, Read, Rerank: Towards End-to-End Multi-Document Reading Comprehension
Minghao Hu, Yuxing Peng, Zhen Huang and Dongsheng Li 2285

Multi-Hop Paragraph Retrieval for Open-Domain Question Answering
Yair Feldman and Ran El-Yaniv 2296

E3: Entailment-driven Extracting and Editing for Conversational Machine Reading
Victor Zhong and Luke Zettlemoyer 2310

Generating Question-Answer Hierarchies
Kalpesh Krishna and Mohit Iyyer 2321

Answering while Summarizing: Multi-task Learning for Multi-hop QA with Evidence Extraction
Kosuke Nishida, Kyosuke Nishida, Masaaki Nagata, Atsushi Otsuka, Itsumi Saito, Hisako Asano and Junji Tomita 2335

Enhancing Pre-Trained Language Representations with Rich Knowledge for Machine Reading Comprehension
An Yang, Quan Wang, Jing Liu, Kai Liu, Yajuan Lyu, Hua Wu, Qiaoqiao She and Sujian Li . . 2346

XQA: A Cross-lingual Open-domain Question Answering Dataset
Jiahua Liu, Yankai Lin, Zhiyuan Liu and Maosong Sun 2358

Compound Probabilistic Context-Free Grammars for Grammar Induction
Yoon Kim, Chris Dyer and Alexander Rush 2369

Semi-supervised Domain Adaptation for Dependency Parsing
Zhenghua Li, Xue Peng, Min Zhang, Rui Wang and Luo Si 2386

Head-Driven Phrase Structure Grammar Parsing on Penn Treebank
Junru Zhou and Hai Zhao 2396

Distantly Supervised Named Entity Recognition using Positive-Unlabeled Learning
Minlong Peng, Xiaoyu Xing, Qi Zhang, Jinlan Fu and Xuanjing Huang 2409

Multi-Task Semantic Dependency Parsing with Policy Gradient for Learning Easy-First Strategies
Shuhei Kurita and Anders Søgaard 2420

GCDT: A Global Context Enhanced Deep Transition Architecture for Sequence Labeling
Yijin Liu, Fandong Meng, Jinchao Zhang, Jinan Xu, Yufeng Chen and Jie Zhou 2431

Unsupervised Learning of PCFGs with Normalizing Flow
Lifeng Jin, Finale Doshi-Velez, Timothy Miller, Lane Schwartz and William Schuler 2442

Variance of Average Surprisal: A Better Predictor for Quality of Grammar from Unsupervised PCFG Induction
Lifeng Jin and William Schuler 2453

Cross-Domain NER using Cross-Domain Language Modeling
Chen Jia, Xiaobo Liang and Yue Zhang 2464

Graph-based Dependency Parsing with Graph Neural Networks
Tao Ji, Yuanbin Wu and Man Lan 2475

Wide-Coverage Neural A Parsing for Minimalist Grammars*
John Torr, Milos Stanojevic, Mark Steedman and Shay B. Cohen 2486

Multi-Modal Sarcasm Detection in Twitter with Hierarchical Fusion Model
Yitao Cai, Huiyu Cai and Xiaojun Wan 2506

Topic-Aware Neural Keyphrase Generation for Social Media Language
Yue Wang, Jing Li, Hou Pong Chan, Irwin King, Michael R. Lyu and Shuming Shi 2516

#YouToo? Detection of Personal Recollections of Sexual Harassment on Social Media
Arijit Ghosh Chowdhury, Ramit Sawhney, Rajiv Ratn Shah and Debanjan Mahata 2527

Multi-task Pairwise Neural Ranking for Hashtag Segmentation
Mounica Maddela, Wei Xu and Daniel Preoţiuc-Pietro 2538

Entity-Centric Contextual Affective Analysis
Anjalie Field and Yulia Tsvetkov 2550

Sentence-Level Evidence Embedding for Claim Verification with Hierarchical Attention Networks
Jing Ma, Wei Gao, Shafiq Joty and Kam-Fai Wong 2561

Predicting Human Activities from User-Generated Content
Steven Wilson and Rada Mihalcea 2572

You Write like You Eat: Stylistic Variation as a Predictor of Social Stratification
Angelo Basile, Albert Gatt and Malvina Nissim 2583

Encoding Social Information with Graph Convolutional Networks forPolitical Perspective Detection in News Media
Chang Li and Dan Goldwasser 2594

Fine-Grained Spoiler Detection from Large-Scale Review Corpora
Mengting Wan, Rishabh Misra, Ndapa Nakashole and Julian McAuley 2605

Celebrity Profiling
Matti Wiegmann, Benno Stein and Martin Potthast 2611

Dataset Creation for Ranking Constructive News Comments
Soichiro Fujita, Hayato Kobayashi and Manabu Okumura 2619

Enhancing Air Quality Prediction with Social Media and Natural Language Processing
Jyun-Yu Jiang, Xue Sun, Wei Wang and Sean Young 2627

Twitter Homophily: Network Based Prediction of User's Occupation
Jiaqi Pan, Rishabh Bhardwaj, Wei Lu, Hai Leong Chieu, Xinghao Pan and Ni Yi Puay 2633

Domain Adaptive Dialog Generation via Meta Learning
Kun Qian and Zhou Yu 2639

Strategies for Structuring Story Generation
Angela Fan, Mike Lewis and Yann Dauphin 2650

Argument Generation with Retrieval, Planning, and Realization
Xinyu Hua, Zhe Hu and Lu Wang 2661

A Simple Recipe towards Reducing Hallucination in Neural Surface Realisation
Feng Nie, Jin-Ge Yao, Jinpeng Wang, Rong Pan and Chin-Yew Lin 2673

Cross-Modal Commentator: Automatic Machine Commenting Based on Cross-Modal Information
Pengcheng Yang, Zhihan Zhang, Fuli Luo, Lei Li, Chengyang Huang and Xu Sun 2680

A Working Memory Model for Task-oriented Dialog Response Generation
Xiuyi Chen, Jiaming Xu and Bo Xu 2687

Cognitive Graph for Multi-Hop Reading Comprehension at Scale
Ming Ding, Chang Zhou, Qibin Chen, Hongxia Yang and Jie Tang 2694

Multi-hop Reading Comprehension across Multiple Documents by Reasoning over Heterogeneous Graphs
Ming Tu, Guangtao Wang, Jing Huang, Yun Tang, Xiaodong He and Bowen Zhou 2704

Explore, Propose, and Assemble: An Interpretable Model for Multi-Hop Reading Comprehension
Yichen Jiang, Nitish Joshi, Yen-Chun Chen and Mohit Bansal 2714

Avoiding Reasoning Shortcuts: Adversarial Evaluation, Training, and Model Development for Multi-Hop QA
Yichen Jiang and Mohit Bansal 2726

Exploiting Explicit Paths for Multi-hop Reading Comprehension
Souvik Kundu, Tushar Khot, Ashish Sabharwal and Peter Clark 2737

Sentence Mover's Similarity: Automatic Evaluation for Multi-Sentence Texts
Elizabeth Clark, Asli Celikyilmaz and Noah A. Smith 2748

Analysis of Automatic Annotation Suggestions for Hard Discourse-Level Tasks in Expert Domains
Claudia Schulz, Christian M. Meyer, Jan Kiesewetter, Michael Sailer, Elisabeth Bauer, Martin R. Fischer, Frank Fischer and Iryna Gurevych 2761

Deep Dominance - How to Properly Compare Deep Neural Models
Rotem Dror, Segev Shlomov and Roi Reichart 2773

We Need to Talk about Standard Splits
Kyle Gorman and Steven Bedrick 2786

Aiming beyond the Obvious: Identifying Non-Obvious Cases in Semantic Similarity Datasets
Nicole Peinelt, Maria Liakata and Dong Nguyen 2792

Putting Evaluation in Context: Contextual Embeddings Improve Machine Translation Evaluation
Nitika Mathur, Timothy Baldwin and Trevor Cohn 2799

Joint Effects of Context and User History for Predicting Online Conversation Re-entries
Xingshan Zeng, Jing Li, Lu Wang and Kam-Fai Wong 2809

CONAN - COunter NArratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech
Yi-Ling Chung, Elizaveta Kuzmenko, Serra Sinem Tekiroglu and Marco Guerini 2819

Categorizing and Inferring the Relationship between the Text and Image of Twitter Posts
Alakananda Vempala and Daniel Preoţiuc-Pietro 2830

Who Sides with Whom? Towards Computational Construction of Discourse Networks for Political Debates
Sebastian Padó, Andre Blessing, Nico Blokker, Erenay Dayanik, Sebastian Haunss and Jonas Kuhn 2841

Analyzing Linguistic Differences between Owner and Staff Attributed Tweets
Daniel Preoţiuc-Pietro and Rita Devlin Marier 2848

Exploring Author Context for Detecting Intended vs Perceived Sarcasm
Silviu Oprea and Walid Magdy 2854

Open Domain Event Extraction Using Neural Latent Variable Models
Xiao Liu, Heyan Huang and Yue Zhang 2860

Multi-Level Matching and Aggregation Network for Few-Shot Relation Classification
Zhi-Xiu Ye and Zhen-Hua Ling 2872

Quantifying Similarity between Relations with Fact Distribution
Weize Chen, Hao Zhu, Xu Han, Zhiyuan Liu and Maosong Sun 2882

Matching the Blanks: Distributional Similarity for Relation Learning
Livio Baldini Soares, Nicholas FitzGerald, Jeffrey Ling and Tom Kwiatkowski 2895

Fine-Grained Temporal Relation Extraction
Siddharth Vashishtha, Benjamin Van Durme and Aaron Steven White 2906

FIESTA: Fast IdEntification of State-of-The-Art models using adaptive bandit algorithms
Henry Moss, Andrew Moore, David Leslie and Paul Rayson 2920

Is Attention Interpretable?
Sofia Serrano and Noah A. Smith 2931

Correlating Neural and Symbolic Representations of Language
Grzegorz Chrupała and Afra Alishahi 2952

Interpretable Neural Predictions with Differentiable Binary Variables
Joost Bastings, Wilker Aziz and Ivan Titov 2963

Transformer-XL: Attentive Language Models beyond a Fixed-Length Context
Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc Le and Ruslan Salakhutdinov 2978

Domain Adaptation of Neural Machine Translation by Lexicon Induction
Junjie Hu, Mengzhou Xia, Graham Neubig and Jaime Carbonell 2989

Reference Network for Neural Machine Translation
Han Fu, Chenghao Liu and Jianling Sun 3002

Retrieving Sequential Information for Non-Autoregressive Neural Machine Translation
Chenze Shao, Yang Feng, Jinchao Zhang, Fandong Meng, Xilin Chen and Jie Zhou 3013

STACL: Simultaneous Translation with Implicit Anticipation and Controllable Latency using Prefix-to-Prefix Framework
Mingbo Ma, Liang Huang, Hao Xiong, Renjie Zheng, Kaibo Liu, Baigong Zheng, Chuanqiang Zhang, Zhongjun He, Hairong Liu, Xing Li, Hua Wu and Haifeng Wang 3025

Look Harder: A Neural Machine Translation Model with Hard Attention
Sathish Reddy Indurthi, Insoo Chung and Sangha Kim 3037

Robust Neural Machine Translation with Joint Textual and Phonetic Embedding
Hairong Liu, Mingbo Ma, Liang Huang, Hao Xiong and Zhongjun He 3044

A Simple and Effective Approach to Automatic Post-Editing with Transfer Learning
Gonçalo M. Correia and André F. T. Martins 3050

Translating Translationese: A Two-Step Approach to Unsupervised Machine Translation
Nima Pourdamghani, Nada Aldarrab, Marjan Ghazvininejad, Kevin Knight and Jonathan May 3057

Training Neural Machine Translation to Apply Terminology Constraints
Georgiana Dinu, Prashant Mathur, Marcello Federico and Yaser Al-Onaizan 3063

Leveraging Local and Global Patterns for Self-Attention Networks
Mingzhou Xu, Derek F. Wong, Baosong Yang, Yue Zhang and Lidia S. Chao 3069

Sentence-Level Agreement for Neural Machine Translation
Mingming Yang, Rui Wang, Kehai Chen, Masao Utiyama, Eiichiro Sumita, Min Zhang and Tiejun Zhao 3076

Multilingual Unsupervised NMT using Shared Encoder and Language-Specific Decoders
Sukanta Sen, Kamal Kumar Gupta, Asif Ekbal and Pushpak Bhattacharyya 3083

Lattice-Based Transformer Encoder for Neural Machine Translation
Fengshun Xiao, Jiangtong Li, Hai Zhao, Rui Wang and Kehai Chen 3090

Multi-Source Cross-Lingual Model Transfer: Learning What to Share
Xilun Chen, Ahmed Hassan Awadallah, Hany Hassan, Wei Wang and Claire Cardie 3098

Unsupervised Multilingual Word Embedding with Limited Resources using Neural Language Models
Takashi Wada, Tomoharu Iwata and Yuji Matsumoto 3113

Choosing Transfer Languages for Cross-Lingual Learning
Yu-Hsiang Lin, Chian-Yu Chen, Jean Lee, Zirui Li, Yuyan Zhang, Mengzhou Xia, Shruti Rijhwani, Junxian He, Zhisong Zhang, Xuezhe Ma, Antonios Anastasopoulos, Patrick Littell and Graham Neubig 3125

CogNet: A Large-Scale Cognate Database
Khuyagbaatar Batsuren, Gabor Bella and Fausto Giunchiglia 3136

Neural Decipherment via Minimum-Cost Flow: From Ugaritic to Linear B
Jiaming Luo, Yuan Cao and Regina Barzilay 3146

Cross-lingual Knowledge Graph Alignment via Graph Matching Neural Network
Kun Xu, Liwei Wang, Mo Yu, Yansong Feng, Yan Song, Zhiguo Wang and Dong Yu 3156

Zero-Shot Cross-Lingual Abstractive Sentence Summarization through Teaching Generation and Attention
Xiangyu Duan, Mingming Yin, Min Zhang, Boxing Chen and Weihua Luo 3162

Improving Low-Resource Cross-lingual Document Retrieval by Reranking with Deep Bilingual Representations
Rui Zhang, Caitlin Westerfield, Sungrok Shim, Garrett Bingham, Alexander Fabbri, William Hu, Neha Verma and Dragomir Radev 3173

Are Girls Neko or Shōjo? Cross-Lingual Alignment of Non-Isomorphic Embeddings with Iterative Normalization
Mozhi Zhang, Keyulu Xu, Ken-ichi Kawarabayashi, Stefanie Jegelka and Jordan Boyd-Graber 3180

MAAM: A Morphology-Aware Alignment Model for Unsupervised Bilingual Lexicon Induction
Pengcheng Yang, Fuli Luo, Peng Chen, Tianyu Liu and Xu Sun 3190

Margin-based Parallel Corpus Mining with Multilingual Sentence Embeddings
Mikel Artetxe and Holger Schwenk 3197

JW300: A Wide-Coverage Parallel Corpus for Low-Resource Languages
Željko Agić and Ivan Vulić 3204

Cross-Lingual Syntactic Transfer through Unsupervised Adaptation of Invertible Projections
Junxian He, Zhisong Zhang, Taylor Berg-Kirkpatrick and Graham Neubig 3211

Unsupervised Joint Training of Bilingual Word Embeddings
Benjamin Marie and Atsushi Fujita 3224

Inferring Concept Hierarchies from Text Corpora via Hyperbolic Embeddings
Matthew Le, Stephen Roller, Laetitia Papaxanthos, Douwe Kiela and Maximilian Nickel 3231

Is Word Segmentation Necessary for Deep Learning of Chinese Representations?
Xiaoya Li, Yuxian Meng, Xiaofei Sun, Qinghong Han, Arianna Yuan and Jiwei Li 3242

Towards Understanding Linear Word Analogies
Kawin Ethayarajh, David Duvenaud and Graeme Hirst 3253

On the Compositionality Prediction of Noun Phrases using Poincaré Embeddings
Abhik Jana, Dima Puzyrev, Alexander Panchenko, Pawan Goyal, Chris Biemann and Animesh Mukherjee 3263

Robust Representation Learning of Biomedical Names
Minh C. Phan, Aixin Sun and Yi Tay 3275

Relational Word Embeddings
Jose Camacho-Collados, Luis Espinosa Anke and Steven Schockaert 3286

Unraveling Antonym's Word Vectors through a Siamese-like Network
Mathias Etcheverry and Dina Wonsever 3297

Incorporating Syntactic and Semantic Information in Word Embeddings using Graph Convolutional Networks
Shikhar Vashishth, Manik Bhandari, Prateek Yadav, Piyush Rai, Chiranjib Bhattacharyya and Partha Talukdar 3308

Word and Document Embedding with vMF-Mixture Priors on Context Word Vectors
Shoaib Jameel and Steven Schockaert 3319

Delta Embedding Learning
Xiao Zhang, Ji Wu and Dejing Dou 3329

Annotation and Automatic Classification of Aspectual Categories
Markus Egg, Helena Prepens and Will Roberts 3335

Putting Words in Context: LSTM Language Models and Lexical Ambiguity
Laura Aina, Kristina Gulordava and Gemma Boleda 3342

Making Fast Graph-based Algorithms with Graph Metric Embeddings
Andrey Kutuzov, Mohammad Dorgham, Oleksiy Oliynyk, Chris Biemann and Alexander Panchenko 3349

Embedding Imputation with Grounded Language Information
Ziyi Yang, Chenguang Zhu, Vin Sachidananda and Eric Darve 3356

The Effectiveness of Simple Hybrid Systems for Hypernym Discovery
William Held and Nizar Habash 3362

BERT-based Lexical Substitution
Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei and Ming Zhou 3368

Exploring Numeracy in Word Embeddings
Aakanksha Naik, Abhilasha Ravichander, Carolyn Rose and Eduard Hovy 3374

HighRES: Highlight-based Reference-less Evaluation of Summarization
Hardy Hardy, Shashi Narayan and Andreas Vlachos 3381

EditNTS: An Neural Programmer-Interpreter Model for Sentence Simplification through Explicit Editing
Yue Dong, Zichao Li, Mehdi Rezagholizadeh and Jackie Chi Kit Cheung 3393

Decomposable Neural Paraphrase Generation
Zichao Li, Xin Jiang, Lifeng Shang and Qun Liu 3403

Transforming Complex Sentences into a Semantic Hierarchy
Christina Niklaus, Matthias Cetto, André Freitas and Siegfried Handschuh 3415

Right for the Wrong Reasons: Diagnosing Syntactic Heuristics in Natural Language Inference
Tom McCoy, Ellie Pavlick and Tal Linzen 3428

Zero-Shot Entity Linking by Reading Entity Descriptions
Lajanugen Logeswaran, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, Jacob Devlin and Honglak Lee 3449

Dual Adversarial Neural Transfer for Low-Resource Named Entity Recognition
Joey Tianyi Zhou, Hao Zhang, Di Jin, Hongyuan Zhu, Meng Fang, Rick Siow Mong Goh and Kenneth Kwok 3461

Scalable Syntax-Aware Language Models Using Knowledge Distillation
Adhiguna Kuncoro, Chris Dyer, Laura Rimell, Stephen Clark and Phil Blunsom 3472

An Imitation Learning Approach to Unsupervised Parsing
Bowen Li, Lili Mou and Frank Keller 3485

Women's Syntactic Resilience and Men's Grammatical Luck: Gender-Bias in Part-of-Speech Tagging and Dependency Parsing
Aparna Garimella, Carmen Banea, Dirk Hovy and Rada Mihalcea 3493

Multilingual Constituency Parsing with Self-Attention and Pre-Training
Nikita Kitaev, Steven Cao and Dan Klein 3499

A Multilingual BPE Embedding Space for Universal Sentiment Lexicon Induction
Mengjie Zhao and Hinrich Schütze 3506

Tree Communication Models for Sentiment Analysis
Yuan Zhang and Yue Zhang 3518

Improved Sentiment Detection via Label Transfer from Monolingual to Synthetic Code-Switched Text
Bidisha Samanta, Niloy Ganguly and Soumen Chakrabarti 3528

Exploring Sequence-to-Sequence Learning in Aspect Term Extraction
Dehong Ma, Sujian Li, Fangzhao Wu, Xing Xie and Houfeng Wang 3538

Aspect Sentiment Classification Towards Question-Answering with Reinforced Bidirectional Attention Network
Jingjing Wang, Changlong Sun, Shoushan Li, Xiaozhong Liu, Luo Si, Min Zhang and Guodong Zhou 3548

ELI5: Long Form Question Answering
Angela Fan, Yacine Jernite, Ethan Perez, David Grangier, Jason Weston and Michael Auli . . . 3558

Textbook Question Answering with Multi-modal Context Graph Understanding and Self-supervised Open-set Comprehension
Daesik Kim, Seonhoon Kim and Nojun Kwak 3568

Generating Question Relevant Captions to Aid Visual Question Answering
Jialin Wu, Zeyuan Hu and Raymond Mooney 3585

Multi-grained Attention with Object-level Grounding for Visual Question Answering
Pingping Huang, Jianhui Huang, Yuqing Guo, Min Qiao and Yong Zhu 3595

Psycholinguistics Meets Continual Learning: Measuring Catastrophic Forgetting in Visual Question Answering
Claudio Greco, Barbara Plank, Raquel Fernández and Raffaella Bernardi 3601

Improving Visual Question Answering by Referring to Generated Paragraph Captions
Hyounghun Kim and Mohit Bansal 3606

Shared-Private Bilingual Word Embeddings for Neural Machine Translation
Xuebo Liu, Derek F. Wong, Yang Liu, Lidia S. Chao, Tong Xiao and Jingbo Zhu 3613

Literary Event Detection
Matthew Sims, Jong Ho Park and David Bamman 3623

Assessing the Ability of Self-Attention Networks to Learn Word Order
Baosong Yang, Longyue Wang, Derek F. Wong, Lidia S. Chao and Zhaopeng Tu 3635

Energy and Policy Considerations for Deep Learning in NLP
Emma Strubell, Ananya Ganesh and Andrew McCallum 3645

What Does BERT Learn about the Structure of Language?
Ganesh Jawahar, Benoît Sagot and Djamé Seddah 3651

A Just and Comprehensive Strategy for Using NLP to Address Online Abuse
David Jurgens, Libby Hemphill and Eshwar Chandrasekharan 3658

Learning from Dialogue after Deployment: Feed Yourself, Chatbot!
Braden Hancock, Antoine Bordes, Pierre-Emmanuel Mazare and Jason Weston 3667

Generating Responses with a Specific Emotion in Dialog
Zhenqiao Song, Xiaoqing Zheng, Lu Liu, Mu Xu and Xuanjing Huang 3685

Semantically Conditioned Dialog Response Generation via Hierarchical Disentangled Self-Attention
Wenhu Chen, Jianshu Chen, Pengda Qin, Xifeng Yan and William Yang Wang 3696

Incremental Learning from Scratch for Task-Oriented Dialogue Systems
Weikang Wang, Jiajun Zhang, Qian Li, Mei-Yuh Hwang, Chengqing Zong and Zhifei Li 3710

ReCoSa: Detecting the Relevant Contexts with Self-Attention for Multi-turn Dialogue Generation
Hainan Zhang, Yanyan Lan, Liang Pang, Jiafeng Guo and Xueqi Cheng 3721

Dialogue Natural Language Inference
Sean Welleck, Jason Weston, Arthur Szlam and Kyunghyun Cho 3731

Budgeted Policy Learning for Task-Oriented Dialogue Systems
Zhirui Zhang, Xiujun Li, Jianfeng Gao and Enhong Chen 3742

Comparison of Diverse Decoding Methods from Conditional Language Models
Daphne Ippolito, Reno Kriz, Joao Sedoc, Maria Kustikova and Chris Callison-Burch 3752

Retrieval-Enhanced Adversarial Training for Neural Response Generation
Qingfu Zhu, Lei Cui, Wei-Nan Zhang, Furu Wei and Ting Liu 3763

Vocabulary Pyramid Network: Multi-Pass Encoding and Decoding with Multi-Level Vocabularies for Response Generation
Cao Liu, Shizhu He, Kang Liu and Jun Zhao 3774

On-device Structured and Context Partitioned Projection Networks
Sujith Ravi and Zornitsa Kozareva 3784

Proactive Human-Machine Conversation with Explicit Conversation Goal
Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian and Haifeng Wang 3794

Learning a Matching Model with Co-teaching for Multi-turn Response Selection in Retrieval-based Dialogue Systems
Jiazhan Feng, Chongyang Tao, Wei Wu, Yansong Feng, Dongyan Zhao and Rui Yan 3805

Learning to Abstract for Memory-augmented Conversational Response Generation
Zhiliang Tian, Wei Bi, Xiaopeng Li and Nevin L. Zhang 3816

Are Training Samples Correlated? Learning to Generate Dialogue Responses with Multiple References
Lisong Qiu, Juntao Li, Wei Bi, Dongyan Zhao and Rui Yan 3826

Pretraining Methods for Dialog Context Representation Learning
Shikib Mehri, Evgeniia Razumovskaia, Tiancheng Zhao and Maxine Eskenazi 3836

A Large-Scale Corpus for Conversation Disentanglement
Jonathan K. Kummerfeld, Sai R. Gouravajhala, Joseph J. Peper, Vignesh Athreya, Chulaka Gunasekara, Jatin Ganhotra, Siva Sankalp Patel, Lazaros C Polymenakos and Walter Lasecki 3846

Self-Supervised Dialogue Learning
Jiawei Wu, Xin Wang and William Yang Wang 3857

Are we there yet? Encoder-decoder neural networks as cognitive models of English past tense inflection
Maria Corkery, Yevgen Matusevych and Sharon Goldwater 3868

A Spreading Activation Framework for Tracking Conceptual Complexity of Texts
Ioana Hulpuș, Sanja Štajner and Heiner Stuckenschmidt 3878

End-to-End Sequential Metaphor Identification Inspired by Linguistic Theories
Rui Mao, Chenghua Lin and Frank Guerin 3888

Diachronic Sense Modeling with Deep Contextualized Word Embeddings: An Ecological View
Renfen Hu, Shen Li and Shichen Liang 3899

Miss Tools and Mr Fruit: Emergent Communication in Agents Learning about Object Affordances
Diane Bouchacourt and Marco Baroni 3909

CNNs found to jump around more skillfully than RNNs: Compositional Generalization in Seq2seq Convolutional Networks
Roberto Dessì and Marco Baroni 3919

Uncovering Probabilistic Implications in Typological Knowledge Bases
Johannes Bjerva, Yova Kementchedjhieva, Ryan Cotterell and Isabelle Augenstein 3924

Is Word Segmentation Child's Play in All Languages?
Georgia R. Loukatou, Steven Moran, Damian Blasi, Sabine Stoll and Alejandrina Cristia 3931

On the Distribution of Deep Clausal Embeddings: A Large Cross-linguistic Study
Damian Blasi, Ryan Cotterell, Lawrence Wolf-Sonkin, Sabine Stoll, Balthasar Bickel and Marco Baroni 3938

Attention-based Conditioning Methods for External Knowledge Integration
Katerina Margatina, Christos Baziotis and Alexandros Potamianos 3944

The KnowRef Coreference Corpus: Removing Gender and Number Cues for Difficult Pronominal Anaphora Resolution
Ali Emami, Paul Trichelair, Adam Trischler, Kaheer Suleman, Hannes Schulz and Jackie Chi Kit Cheung 3952

StRE: Self Attentive Edit Quality Prediction in Wikipedia
Soumya Sarkar, Bhanu Prakash Reddy, Sandipan Sikdar and Animesh Mukherjee 3962

How Large Are Lions? Inducing Distributions over Quantitative Attributes
Yanai Elazar, Abhijit Mahabal, Deepak Ramachandran, Tania Bedrax-Weiss and Dan Roth . . 3973

Fine-Grained Sentence Functions for Short-Text Conversation
Wei Bi, Jun Gao, Xiaojiang Liu and Shuming Shi 3984

Give Me More Feedback II: Annotating Thesis Strength and Related Attributes in Student Essays
Zixuan Ke, Hrishikesh Inamdar, Hui Lin and Vincent Ng 3994

Crowdsourcing and Validating Event-focused Emotion Corpora for German and English
Enrica Troiano, Sebastian Padó and Roman Klinger 4005

Pay Attention when you Pay the Bills. A Multilingual Corpus with Dependency-based and Semantic Annotation of Collocations.
Marcos Garcia, Marcos García Salido, Susana Sotelo, Estela Mosqueira and Margarita Alonso-Ramos 4012

Does it Make Sense? And Why? A Pilot Study for Sense Making and Explanation
Cunxiang Wang, Shuailong Liang, Yue Zhang, Xiaonan Li and Tian Gao 4020

Large Dataset and Language Model Fun-Tuning for Humor Recognition
Vladislav Blinov, Valeria Bolotova-Baranova and Pavel Braslavski 4027

Towards Language Agnostic Universal Representations
Armen Aghajanyan, Xia Song and Saurabh Tiwary 4033

Leveraging Meta Information in Short Text Aggregation
He Zhao, Lan Du, Guanfeng Liu and Wray Buntine 4042

Exploiting Invertible Decoders for Unsupervised Sentence Representation Learning
Shuai Tang and Virginia R. de Sa 4050

Self-Attentive, Multi-Context One-Class Classification for Unsupervised Anomaly Detection on Text
Lukas Ruff, Yury Zemlyanskiy, Robert Vandermeulen, Thomas Schnake and Marius Kloft . . . 4061

Hubless Nearest Neighbor Search for Bilingual Lexicon Induction
Jiaji Huang, Qiang Qiu and Kenneth Church 4072

Distant Learning for Entity Linking with Automatic Noise Detection
Phong Le and Ivan Titov 4081

Learning How to Active Learn by Dreaming
Thuy-Trang Vu, Ming Liu, Dinh Phung and Gholamreza Haffari 4091

Few-Shot Representation Learning for Out-Of-Vocabulary Words
Ziniu Hu, Ting Chen, Kai-Wei Chang and Yizhou Sun 4102

Neural Temporality Adaptation for Document Classification: Diachronic Word Embeddings and Domain Adaptation Models
Xiaolei Huang and Michael J. Paul 4113

Learning Transferable Feature Representations Using Neural Networks
Himanshu Sharad Bhatt, Shourya Roy, Arun Rajkumar and Sriranjani Ramakrishnan 4124

Bayes Test of Precision, Recall, and F1 Measure for Comparison of Two Natural Language Processing Models
Ruibo Wang and Jihong Li 4135

TIGS: An Inference Algorithm for Text Infilling with Gradient Search
Dayiheng Liu, Jie Fu, Pengfei Liu and Jiancheng Lv 4146

Keeping Notes: Conditional Natural Language Generation with a Scratchpad Encoder
Ryan Benmalek, Madian Khabsa, Suma Desu, Claire Cardie and Michele Banko 4157

Using Automatically Extracted Minimum Spans to Disentangle Coreference Evaluation from Boundary Detection
Nafise Sadat Moosavi, Leo Born, Massimo Poesio and Michael Strube 4168

Revisiting Joint Modeling of Cross-document Entity and Event Coreference Resolution
Shany Barhom, Vered Shwartz, Alon Eirew, Michael Bugert, Nils Reimers and Ido Dagan . . . 4179

A Unified Linear-Time Framework for Sentence-Level Discourse Parsing
Xiang Lin, Shafiq Joty, Prathyusha Jwalapuram and M Saiful Bari 4190

Employing the Correspondence of Relations and Connectives to Identify Implicit Discourse Relations via Label Embeddings
Linh The Nguyen, Linh Van Ngo, Khoat Than and Thien Huu Nguyen 4201

Do You Know That Florence Is Packed with Visitors? Evaluating State-of-the-art Models of Speaker Commitment
Nanjiang Jiang and Marie-Catherine de Marneffe 4208

Multi-Relational Script Learning for Discourse Relations
I-Ta Lee and Dan Goldwasser 4214

Open-Domain Why-Question Answering with Adversarial Learning to Encode Answer Texts
Jong-Hoon Oh, Kazuma Kadowaki, Julien Kloetzer, Ryu Iida and Kentaro Torisawa 4227

Learning to Ask Unanswerable Questions for Machine Reading Comprehension
Haichao Zhu, Li Dong, Furu Wei, Wenhui Wang, Bing Qin and Ting Liu 4238

Compositional Questions Do Not Necessitate Multi-hop Reasoning
Sewon Min, Eric Wallace, Sameer Singh, Matt Gardner, Hannaneh Hajishirzi and Luke Zettlemoyer 4249

Improving Question Answering over Incomplete KBs with Knowledge-Aware Reader
Wenhan Xiong, Mo Yu, Shiyu Chang, Xiaoxiao Guo and William Yang Wang 4258

AdaNSP: Uncertainty-driven Adaptive Decoding in Neural Semantic Parsing
Xiang Zhang, Shizhu He, Kang Liu and Jun Zhao 4265

The Language of Legal and Illegal Activity on the Darknet
Leshem Choshen, Dan Eldad, Daniel Hershcovich, Elior Sulem and Omri Abend 4271

Eliciting Knowledge from Experts: Automatic Transcript Parsing for Cognitive Task Analysis
Junyi Du, He Jiang, Jiaming Shen and Xiang Ren 4280

Course Concept Expansion in MOOCs with External Knowledge and Interactive Game
Jifan Yu, Chenyu Wang, Gan Luo, Lei Hou, Juanzi Li, Zhiyuan Liu and Jie Tang 4292

Towards Near-imperceptible Steganographic Text
Falcon Dai and Zheng Cai 4303

Inter-sentence Relation Extraction with Document-level Graph Convolutional Neural Network
Sunil Kumar Sahu, Fenia Christopoulou, Makoto Miwa and Sophia Ananiadou 4309

Neural Legal Judgment Prediction in English
Ilias Chalkidis, Ion Androutsopoulos and Nikolaos Aletras 4317

Robust Neural Machine Translation with Doubly Adversarial Inputs
Yong Cheng, Lu Jiang and Wolfgang Macherey 4324

Bridging the Gap between Training and Inference for Neural Machine Translation
Wen Zhang, Yang Feng, Fandong Meng, Di You and Qun Liu 4334

Beyond BLEU:Training Neural Machine Translation with Semantic Similarity
John Wieting, Taylor Berg-Kirkpatrick, Kevin Gimpel and Graham Neubig 4344

AutoML Strategy Based on Grammatical Evolution: A Case Study about Knowledge Discovery from Text
Suilan Estevez-Velarde, Yoan Gutiérrez, Andrés Montoyo and Yudivián Almeida-Cruz 4356

Distilling Discrimination and Generalization Knowledge for Event Detection via Delta-Representation Learning
Yaojie Lu, Hongyu Lin, Xianpei Han and Le Sun 4366

Chinese Relation Extraction with Multi-Grained Information and External Linguistic Knowledge
Ziran Li, Ning Ding, Zhiyuan Liu, Haitao Zheng and Ying Shen 4377

A2N: Attending to Neighbors for Knowledge Graph Inference
Trapit Bansal, Da-Cheng Juan, Sujith Ravi and Andrew McCallum 4387

Graph based Neural Networks for Event Factuality Prediction using Syntactic and Semantic Structures
Amir Pouran Ben Veyseh, Thien Huu Nguyen and Dejing Dou 4393

Embedding Time Expressions for Deep Temporal Ordering Models
Tanya Goyal and Greg Durrett 4400

Episodic Memory Reader: Learning What to Remember for Question Answering from Streaming Data
Moonsu Han, Minki Kang, Hyunwoo Jung and Sung Ju Hwang 4407

Selection Bias Explorations and Debias Methods for Natural Language Sentence Matching Datasets
Guanhua Zhang, Bing Bai, Jian Liang, Kun Bai, Shiyu Chang, Mo Yu, Conghui Zhu and Tiejun Zhao 4418

Real-Time Open-Domain Question Answering with Dense-Sparse Phrase Index
Minjoon Seo, Jinhyuk Lee, Tom Kwiatkowski, Ankur Parikh, Ali Farhadi and Hannaneh Hajishirzi 4430

Language Modeling with Shared Grammar
Yuyu Zhang and Le Song 4442

Zero-Shot Semantic Parsing for Instructions
Ofer Givoli and Roi Reichart 4454

Can You Tell Me How to Get Past Sesame Street? Sentence-Level Pretraining Beyond Language Modeling
Alex Wang, Jan Hula, Patrick Xia, Raghavendra Pappagari, R. Thomas McCoy, Roma Patel, Najoung Kim, Ian Tenney, Yinghui Huang, Katherin Yu, Shuning Jin, Berlin Chen, Benjamin Van Durme, Edouard Grave, Ellie Pavlick and Samuel R. Bowman 4465

Complex Question Decomposition for Semantic Parsing
Haoyu Zhang, Jingjing Cai, Jianjun Xu and Ji Wang 4477

Multi-Task Deep Neural Networks for Natural Language Understanding
Xiaodong Liu, Pengcheng He, Weizhu Chen and Jianfeng Gao 4487

DisSent: Learning Sentence Representations from Explicit Discourse Relations
Allen Nie, Erin Bennett and Noah Goodman 4497

SParC: Cross-Domain Semantic Parsing in Context
Tao Yu, Rui Zhang, Michihiro Yasunaga, Yi Chern Tan, Xi Victoria Lin, Suyi Li, Heyang Er, Irene Li, Bo Pang, Tao Chen, Emily Ji, Shreya Dixit, David Proctor, Sungrok Shim, Jonathan Kraft, Vincent Zhang, Caiming Xiong, Richard Socher and Dragomir Radev 4511

Towards Complex Text-to-SQL in Cross-Domain Database with Intermediate Representation
Jiaqi Guo, Zecheng Zhan, Yan Gao, Yan Xiao, Jian-Guang Lou, Ting Liu and Dongmei Zhang4524

EigenSent: Spectral sentence embeddings using higher-order Dynamic Mode Decomposition
Subhradeep Kayal and George Tsatsaronis 4536

SemBleu: A Robust Metric for AMR Parsing Evaluation
Linfeng Song and Daniel Gildea 4547

Reranking for Neural Semantic Parsing
Pengcheng Yin and Graham Neubig 4553

Representing Schema Structure with Graph Neural Networks for Text-to-SQL Parsing
Ben Bogin, Jonathan Berant and Matt Gardner 4560

Human vs. Muppet: A Conservative Estimate of Human Performance on the GLUE Benchmark
Nikita Nangia and Samuel R. Bowman 4566

Compositional Semantic Parsing across Graphbanks
Matthias Lindemann, Jonas Groschwitz and Alexander Koller 4576

Rewarding Smatch: Transition-Based AMR Parsing with Reinforcement Learning
Tahira Naseem, Abhishek Shah, Hui Wan, Radu Florian, Salim Roukos and Miguel Ballesteros 4586

BERT Rediscovers the Classical NLP Pipeline
Ian Tenney, Dipanjan Das and Ellie Pavlick 4593

Simple and Effective Paraphrastic Similarity from Parallel Translations
John Wieting, Kevin Gimpel, Graham Neubig and Taylor Berg-Kirkpatrick 4602

Second-Order Semantic Dependency Parsing with End-to-End Neural Networks
Xinyu Wang, Jingxian Huang and Kewei Tu 4609

Towards Multimodal Sarcasm Detection (An _Obviously_ Perfect Paper)
Santiago Castro, Devamanyu Hazarika, Verónica Pérez-Rosas, Roger Zimmermann, Rada Mihalcea and Soujanya Poria 4619

Determining Relative Argument Specificity and Stance for Complex Argumentative Structures
Esin Durmus, Faisal Ladhak and Claire Cardie 4630

Latent Variable Sentiment Grammar
Liwen Zhang, Kewei Tu and Yue Zhang 4642

An Investigation of Transfer Learning-Based Sentiment Analysis in Japanese
Enkhbold Bataa and Joshua Wu 4652

Probing Neural Network Comprehension of Natural Language Arguments
Timothy Niven and Hung-Yu Kao 4658

Recognising Agreement and Disagreement between Stances with Reason Comparing Networks
Chang Xu, Cecile Paris, Surya Nepal and Ross Sparks 4665

Toward Comprehensive Understanding of a Sentiment Based on Human Motives
Naoki Otani and Eduard Hovy 4672

Context-aware Embedding for Targeted Aspect-based Sentiment Analysis
Bin Liang, Jiachen Du, Ruifeng Xu, Binyang Li and Hejiao Huang 4678

Yes, we can! Mining Arguments in 50 Years of US Presidential Campaign Debates
Shohreh Haddadan, Elena Cabrio and Serena Villata 4684

An Empirical Study of Span Representations in Argumentation Structure Parsing
Tatsuki Kuribayashi, Hiroki Ouchi, Naoya Inoue, Paul Reisert, Toshinori Miyoshi, Jun Suzuki and Kentaro Inui 4691

Simple and Effective Text Matching with Richer Alignment Features
Runqi Yang, Jianhai Zhang, Xing Gao, Feng Ji and Haiqing Chen 4699

Learning Attention-based Embeddings for Relation Prediction in Knowledge Graphs
Deepak Nathani, Jatin Chauhan, Charu Sharma and Manohar Kaul 4710

Neural Network Alignment for Sentential Paraphrases
Jessica Ouyang and Kathy McKeown 4724

Duality of Link Prediction and Entailment Graph Induction
Mohammad Javad Hosseini, Shay B. Cohen, Mark Johnson and Mark Steedman 4736

A Cross-Sentence Latent Variable Model for Semi-Supervised Text Sequence Matching
Jihun Choi, Taeuk Kim and Sang-goo Lee 4747

COMET: Commonsense Transformers for Automatic Knowledge Graph Construction
Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz and Yejin Choi 4762

Detecting Subevents using Discourse and Narrative Features
Mohammed Aldawsari and Mark Finlayson 4780

HellaSwag: Can a Machine Really Finish Your Sentence?
Rowan Zellers, Ari Holtzman, Yonatan Bisk, Ali Farhadi and Yejin Choi 4791

Unified Semantic Parsing with Weak Supervision
Priyanka Agrawal, Ayushi Dalmia, Parag Jain, Abhishek Bansal, Ashish Mittal and Karthik Sankaranarayanan 4801

Every Child Should Have Parents: A Taxonomy Refinement Algorithm Based on Hyperbolic Term Embeddings
Rami Aly, Shantanu Acharya, Alexander Ossa, Arne Köhn, Chris Biemann and Alexander Panchenko 4811

Learning to Rank for Plausible Plausibility
Zhongyang Li, Tongfei Chen and Benjamin Van Durme 4818

Generalized Tuning of Distributional Word Vectors for Monolingual and Cross-Lingual Lexical Entailment
Goran Glavaš and Ivan Vulić 4824

Attention Is (not) All You Need for Commonsense Reasoning
Tassilo Klein and Moin Nabi 4831

A Surprisingly Robust Trick for the Winograd Schema Challenge
Vid Kocijan, Ana-Maria Cretu, Oana-Maria Camburu, Yordan Yordanov and Thomas Lukasiewicz 4837

Coherent Comments Generation for Chinese Articles with a Graph-to-Sequence Model
Wei Li, Jingjing Xu, Yancheng He, ShengLi Yan, Yunfang Wu and Xu Sun 4843

Interconnected Question Generation with Coreference Alignment and Conversation Flow Modeling
Yifan Gao, Piji Li, Irwin King and Michael R. Lyu 4853

Cross-Lingual Training for Automatic Question Generation
Vishwajeet Kumar, Nitish Joshi, Arijit Mukherjee, Ganesh Ramakrishnan and Preethi Jyothi . 4863

A Hierarchical Reinforced Sequence Operation Method for Unsupervised Text Style Transfer
Chen Wu, Xuancheng Ren, Fuli Luo and Xu Sun 4873

Handling Divergent Reference Texts when Evaluating Table-to-Text Generation
Bhuwan Dhingra, Manaal Faruqui, Ankur Parikh, Ming-Wei Chang, Dipanjan Das and William Cohen 4884

Unsupervised Question Answering by Cloze Translation
Patrick Lewis, Ludovic Denoyer and Sebastian Riedel 4896

MultiQA: An Empirical Investigation of Generalization and Transfer in Reading Comprehension
Alon Talmor and Jonathan Berant 4911

Simple and Effective Curriculum Pointer-Generator Networks for Reading Comprehension over Long Narratives
Yi Tay, Shuohang Wang, Anh Tuan Luu, Jie Fu, Minh C. Phan, Xingdi Yuan, Jinfeng Rao, Siu Cheung Hui and Aston Zhang 4922

Explain Yourself! Leveraging Language Models for Commonsense Reasoning
Nazneen Fatema Rajani, Bryan McCann, Caiming Xiong and Richard Socher 4932

Interpretable Question Answering on Knowledge Bases and Text
Alona Sydorova, Nina Poerner and Benjamin Roth 4943

A Resource-Free Evaluation Metric for Cross-Lingual Word Embeddings Based on Graph Modularity
Yoshinari Fujinuma, Jordan Boyd-Graber and Michael J. Paul 4952

Multilingual and Cross-Lingual Graded Lexical Entailment
Ivan Vulić, Simone Paolo Ponzetto and Goran Glavaš 4963

What Kind of Language Is Hard to Language-Model?
Sebastian J. Mielke, Ryan Cotterell, Kyle Gorman, Brian Roark and Jason Eisner 4975

Analyzing the Limitations of Cross-lingual Word Embedding Mappings
Aitor Ormazabal, Mikel Artetxe, Gorka Labaka, Aitor Soroa and Eneko Agirre 4990

How Multilingual is Multilingual BERT?
Telmo Pires, Eva Schlinger and Dan Garrette 4996

Bilingual Lexicon Induction through Unsupervised Machine Translation
Mikel Artetxe, Gorka Labaka and Eneko Agirre 5002

Automatically Identifying Complaints in Social Media
Daniel Preoţiuc-Pietro, Mihaela Gaman and Nikolaos Aletras 5008

TWEETQA: A Social Media Focused Question Answering Dataset
Wenhan Xiong, Jiawei Wu, Hong Wang, Vivek Kulkarni, Mo Yu, Shiyu Chang, Xiaoxiao Guo and William Yang Wang 5020

Asking the Crowd: Question Analysis, Evaluation and Generation for Open Discussion on Online Forums
Zi Chai, Xinyu Xing, Xiaojun Wan and Bo Huang 5032

Tree LSTMs with Convolution Units to Predict Stance and Rumor Veracity in Social Media Conversations
Sumeet Kumar and Kathleen Carley 5047

HIBERT: Document Level Pre-training of Hierarchical Bidirectional Transformers for Document Summarization
Xingxing Zhang, Furu Wei and Ming Zhou 5059

Hierarchical Transformers for Multi-Document Summarization
Yang Liu and Mirella Lapata 5070

Abstractive Text Summarization Based on Deep Learning and Semantic Content Generalization
Panagiotis Kouris, Georgios Alexandridis and Andreas Stafylopatis 5082

Studying Summarization Evaluation Metrics in the Appropriate Scoring Range
Maxime Peyrard 5093

Simple Unsupervised Summarization by Contextual Matching
Jiawei Zhou and Alexander Rush 5101

Generating Summaries with Topic Templates and Structured Convolutional Decoders
Laura Perez-Beltrachini, Yang Liu and Mirella Lapata 5107

Morphological Irregularity Correlates with Frequency
Shijie Wu, Ryan Cotterell and Timothy O'Donnell 5117

Like a Baby: Visually Situated Neural Language Acquisition
Alexander Ororbia, Ankur Mali, Matthew Kelly and David Reitter 5127

Relating Simple Sentence Representations in Deep Neural Networks and the Brain
Sharmistha Jat, Hao Tang, Partha Talukdar and Tom Mitchell 5137

Modeling Affirmative and Negated Action Processing in the Brain with Lexical and Compositional Semantic Models
Vesna Djokic, Jean Maillard, Luana Bulat and Ekaterina Shutova 5155

Word-order Biases in Deep-agent Emergent Communication
Rahma Chaabouni, Eugene Kharitonov, Alessandro Lazaric, Emmanuel Dupoux and Marco Baroni 5166

NNE: A Dataset for Nested Named Entity Recognition in English Newswire
Nicky Ringland, Xiang Dai, Ben Hachey, Sarvnaz Karimi, Cecile Paris and James R. Curran . 5176

Sequence-to-Nuggets: Nested Entity Mention Detection via Anchor-Region Networks
Hongyu Lin, Yaojie Lu, Xianpei Han and Le Sun 5182

Improving Textual Network Embedding with Global Attention via Optimal Transport
Liqun Chen, Guoyin Wang, Chenyang Tao, Dinghan Shen, Pengyu Cheng, Xinyuan Zhang, Wenlin Wang, Yizhe Zhang and Lawrence Carin 5193

Identification of Tasks, Datasets, Evaluation Metrics, and Numeric Scores for Scientific Leaderboards Construction
Yufang Hou, Charles Jochim, Martin Gleize, Francesca Bonin and Debasis Ganguly 5203

Scaling up Open Tagging from Tens to Thousands: Comprehension Empowered Attribute Value Extraction from Product Title
Huimin Xu, Wenting Wang, Xin Mao, Xinyu Jiang and Man Lan 5214

Incorporating Linguistic Constraints into Keyphrase Generation
Jing Zhao and Yuxiang Zhang 5224

A Unified Multi-task Adversarial Learning Framework for Pharmacovigilance Mining
Shweta Yadav, Asif Ekbal, Sriparna Saha and Pushpak Bhattacharyya 5234

Quantity Tagger: A Latent-Variable Sequence Labeling Approach to Solving Addition-Subtraction Word Problems
Yanyan Zou and Wei Lu 5246

A Deep Reinforced Sequence-to-Set Model for Multi-Label Classification
Pengcheng Yang, Fuli Luo, Shuming Ma, Junyang Lin and Xu Sun 5252

Joint Slot Filling and Intent Detection via Capsule Neural Networks
Chenwei Zhang, Yaliang Li, Nan Du, Wei Fan and Philip Yu 5259

Neural Aspect and Opinion Term Extraction with Mined Rules as Weak Supervision
Hongliang Dai and Yangqiu Song 5268

Cost-sensitive Regularization for Label Confusion-aware Event Detection
Hongyu Lin, Yaojie Lu, Xianpei Han and Le Sun 5278

Exploring Pre-trained Language Models for Event Extraction and Generation
Sen Yang, Dawei Feng, Linbo Qiao, Zhigang Kan and Dongsheng Li 5284

Improving Open Information Extraction via Iterative Rank-Aware Learning
Zhengbao Jiang, Pengcheng Yin and Graham Neubig 5295

Towards Improving Neural Named Entity Recognition with Gazetteers
Tianyu Liu, Jin-Ge Yao and Chin-Yew Lin 5301

Span-Level Model for Relation Extraction
Kalpit Dixit and Yaser Al-Onaizan 5308

Enhancing Unsupervised Generative Dependency Parser with Contextual Information
Wenjuan Han, Yong Jiang and Kewei Tu 5315

Neural Architectures for Nested NER through Linearization
Jana Straková, Milan Straka and Jan Hajic 5326

Online Infix Probability Computation for Probabilistic Finite Automata
Marco Cognetta, Yo-Sub Han and Soon Chan Kwon 5332

How to Best Use Syntax in Semantic Role Labelling
Yufei Wang, Mark Johnson, Stephen Wan, Yifang Sun and Wei Wang 5338

PTB Graph Parsing with Tree Approximation
Yoshihide Kato and Shigeki Matsubara 5344

Sequence Labeling Parsing by Learning across Representations
Michalina Strzyz, David Vilares and Carlos Gómez-Rodríguez 5350

A Prism Module for Semantic Disentanglement in Name Entity Recognition
Kun Liu, Shen Li, Daqi Zheng, Zhengdong Lu, Sheng Gao and Si Li 5358

Label-Agnostic Sequence Labeling by Copying Nearest Neighbors
Sam Wiseman and Karl Stratos 5363

Towards Empathetic Open-domain Conversation Models: A New Benchmark and Dataset
Hannah Rashkin, Eric Michael Smith, Margaret Li and Y-Lan Boureau 5370

Know More about Each Other: Evolving Dialogue Strategy via Compound Assessment
Siqi Bao, Huang He, Fan Wang, Rongzhong Lian and Hua Wu 5382

Training Neural Response Selection for Task-Oriented Dialogue Systems
Matthew Henderson, Ivan Vulić, Daniela Gerz, Iñigo Casanueva, Paweł Budzianowski, Sam Coope, Georgios Spithourakis, Tsung-Hsien Wen, Nikola Mrkšić and Pei-Hao Su 5392

Collaborative Dialogue in Minecraft
Anjali Narayan-Chen, Prashant Jayannavar and Julia Hockenmaier 5405

Neural Response Generation with Meta-words
Can Xu, Wei Wu, Chongyang Tao, Huang Hu, Matt Schuerman and Ying Wang 5416

Conversing by Reading: Contentful Neural Conversation with On-demand Machine Reading
Lianhui Qin, Michel Galley, Chris Brockett, Xiaodong Liu, Xiang Gao, Bill Dolan, Yejin Choi and Jianfeng Gao 5427

Ordinal and Attribute Aware Response Generation in a Multimodal Dialogue System
Hardik Chauhan, Mauajama Firdaus, Asif Ekbal and Pushpak Bhattacharyya 5437

Memory Consolidation for Contextual Spoken Language Understanding with Dialogue Logistic Inference
He Bai, Yu Zhou, Jiajun Zhang and Chengqing Zong 5448

Personalizing Dialogue Agents via Meta-Learning
Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu and Pascale Fung 5454

Reading Turn by Turn: Hierarchical Attention Architecture for Spoken Dialogue Comprehension
Zhengyuan Liu and Nancy Chen 5460

A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling
Haihong E, Peiqing Niu, Zhongfu Chen and Meina Song 5467

Dual Supervised Learning for Natural Language Understanding and Generation
Shang-Yu Su, Chao-Wei Huang and Yun-Nung Chen 5472

SUMBT: Slot-Utterance Matching for Universal and Scalable Belief Tracking
Hwaran Lee, Jinsik Lee and Tae-Yoon Kim 5478

Robust Zero-Shot Cross-Domain Slot Filling with Example Values
Darsh Shah, Raghav Gupta, Amir Fayazi and Dilek Hakkani-Tur 5484

Deep Unknown Intent Detection with Margin Loss
Ting-En Lin and Hua Xu 5491

Modeling Semantic Relationship in Multi-turn Conversations with Hierarchical Latent Variables
Lei Shen, Yang Feng and Haolan Zhan 5497

Rationally Reappraising ATIS-based Dialogue Systems
Jingcheng Niu and Gerald Penn 5503

Learning Latent Trees with Stochastic Perturbations and Differentiable Dynamic Programming
Caio Corro and Ivan Titov 5508

Neural-based Chinese Idiom Recommendation for Enhancing Elegance in Essay Writing
Yuanchao Liu, Bo Pang and Bingquan Liu 5522

Better Exploiting Latent Variables in Text Modeling
Canasai Kruengkrai 5527

Misleading Failures of Partial-input Baselines
Shi Feng, Eric Wallace and Jordan Boyd-Graber 5533

Soft Contextual Data Augmentation for Neural Machine Translation
Fei Gao, Jinhua Zhu, Lijun Wu, Yingce Xia, Tao Qin, Xueqi Cheng, Wengang Zhou and Tie-Yan Liu 5539

Reversing Gradients in Adversarial Domain Adaptation for Question Deduplication and Textual Entailment Tasks
Anush Kamath, Sparsh Gupta and Vitor Carvalho 5545

Towards Integration of Statistical Hypothesis Tests into Deep Neural Networks
Ahmad Aghaebrahimian and Mark Cieliebak 5551

Depth Growing for Neural Machine Translation
Lijun Wu, Yiren Wang, Yingce Xia, Fei Tian, Fei Gao, Tao Qin, Jianhuang Lai and Tie-Yan Liu 5558

Generating Fluent Adversarial Examples for Natural Languages
Huangzhao Zhang, Hao Zhou, Ning Miao and Lei Li 5564

Towards Explainable NLP: A Generative Explanation Framework for Text Classification
Hui Liu, Qingyu Yin and William Yang Wang 5570

Combating Adversarial Misspellings with Robust Word Recognition
Danish Pruthi, Bhuwan Dhingra and Zachary C. Lipton 5582

An Empirical Investigation of Structured Output Modeling for Graph-based Neural Dependency Parsing
Zhisong Zhang, Xuezhe Ma and Eduard Hovy 5592

Observing Dialogue in Therapy: Categorizing and Forecasting Behavioral Codes
Jie Cao, Michael Tanana, Zac Imel, Eric Poitras, David Atkins and Vivek Srikumar 5599

Multimodal Transformer Networks for End-to-End Video-Grounded Dialogue Systems
Hung Le, Doyen Sahoo, Nancy Chen and Steven Hoi 5612

Target-Guided Open-Domain Conversation
Jianheng Tang, Tiancheng Zhao, Chenyan Xiong, Xiaodan Liang, Eric Xing and Zhiting Hu . 5624

Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good
Xuewei Wang, Weiyan Shi, Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang and Zhou Yu 5635

Improving Neural Conversational Models with Entropy-Based Data Filtering
Richárd Csáky, Patrik Purgai and Gábor Recski 5650

Zero-shot Word Sense Disambiguation using Sense Definition Embeddings
Sawan Kumar, Sharmistha Jat, Karan Saxena and Partha Talukdar 5670

Language Modelling Makes Sense: Propagating Representations through WordNet for Full-Coverage Word Sense Disambiguation
Daniel Loureiro and Alípio Jorge 5682

Word2Sense: Sparse Interpretable Word Embeddings
Abhishek Panigrahi, Harsha Vardhan Simhadri and Chiranjib Bhattacharyya 5692

Modeling Semantic Compositionality with Sememe Knowledge
Fanchao Qi, Junjie Huang, Chenghao Yang, Zhiyuan Liu, Xiao Chen, Qun Liu and Maosong Sun 5706

Predicting Humorousness and Metaphor Novelty with Gaussian Process Preference Learning
Edwin Simpson, Erik-Lân Do Dinh, Tristan Miller and Iryna Gurevych 5716

Empirical Linguistic Study of Sentence Embeddings
Katarzyna Krasnowska-Kieraś and Alina Wróblewska 5729

Probing for Semantic Classes: Diagnosing the Meaning Content of Word Embeddings
Yadollah Yaghoobzadeh, Katharina Kann, T. J. Hazen, Eneko Agirre and Hinrich Schütze . . . 5740

Deep Neural Model Inspection and Comparison via Functional Neuron Pathways
James Fiacco, Samridhi Choudhary and Carolyn Rose 5754

Collocation Classification with Unsupervised Relation Vectors
Luis Espinosa Anke, Steven Schockaert and Leo Wanner 5765

Corpus-based Check-up for Thesaurus
Natalia Loukachevitch 5773

Confusionset-guided Pointer Networks for Chinese Spelling Check
Dingmin Wang, Yi Tay and Li Zhong 5780

Generalized Data Augmentation for Low-Resource Translation
Mengzhou Xia, Xiang Kong, Antonios Anastasopoulos and Graham Neubig 5786

Analyzing Multi-Head Self-Attention: Specialized Heads Do the Heavy Lifting, the Rest Can Be Pruned
Elena Voita, David Talbot, Fedor Moiseev, Rico Sennrich and Ivan Titov 5797

Better OOV Translation with Bilingual Terminology Mining
Matthias Huck, Viktor Hangya and Alexander Fraser 5809

Simultaneous Translation with Flexible Policy via Restricted Imitation Learning
Baigong Zheng, Renjie Zheng, Mingbo Ma and Liang Huang 5816

Target Conditioned Sampling: Optimizing Data Selection for Multilingual Neural Machine Translation
Xinyi Wang and Graham Neubig 5823

Adversarial Learning of Privacy-Preserving Text Representations for De-Identification of Medical Records
Max Friedrich, Arne Köhn, Gregor Wiedemann and Chris Biemann 5829

Merge and Label: A Novel Neural Network Architecture for Nested NER
Joseph Fisher and Andreas Vlachos 5840

Low-resource Deep Entity Resolution with Transfer and Active Learning
Jungo Kasai, Kun Qian, Sairam Gurajada, Yunyao Li and Lucian Popa 5851

A Semi-Markov Structured Support Vector Machine Model for High-Precision Named Entity Recognition
Ravneet Arora, Chen-Tse Tsai, Ketevan Tsereteli, Prabhanjan Kambadur and Yi Yang 5862

Using Human Attention to Extract Keyphrase from Microblog Post
Yingyi Zhang and Chengzhi Zhang 5867

Model-Agnostic Meta-Learning for Relation Classification with Limited Supervision
Abiola Obamuyide and Andreas Vlachos 5873

Variational Pretraining for Semi-supervised Text Classification
Suchin Gururangan, Tam Dang, Dallas Card and Noah A. Smith 5880

Task Refinement Learning for Improved Accuracy and Stability of Unsupervised Domain Adaptation
Yftah Ziser and Roi Reichart 5895

Optimal Transport-based Alignment of Learned Character Representations for String Similarity
Derek Tam, Nicholas Monath, Ari Kobren, Aaron Traylor, Rajarshi Das and Andrew McCallum 5907

The Referential Reader: A Recurrent Entity Network for Anaphora Resolution
Fei Liu, Luke Zettlemoyer and Jacob Eisenstein 5918

Interpolated Spectral NGram Language Models
Ariadna Quattoni and Xavier Carreras 5926

BAM! Born-Again Multi-Task Networks for Natural Language Understanding
Kevin Clark, Minh-Thang Luong, Urvashi Khandelwal, Christopher D. Manning and Quoc V. Le 5931

Curate and Generate: A Corpus and Method for Joint Control of Semantics and Style in Neural NLG
Shereen Oraby, Vrindavan Harrison, Abteen Ebrahimi and Marilyn Walker 5938

Automated Chess Commentator Powered by Neural Chess Engine
Hongyu Zang, Zhiwei Yu and Xiaojun Wan 5952

Barack's Wife Hillary: Using Knowledge Graphs for Fact-Aware Language Modeling
Robert Logan, Nelson F. Liu, Matthew E. Peters, Matt Gardner and Sameer Singh 5962

Controllable Paraphrase Generation with a Syntactic Exemplar
Mingda Chen, Qingming Tang, Sam Wiseman and Kevin Gimpel 5972

Towards Comprehensive Description Generation from Factual Attribute-value Tables
Tianyu Liu, Fuli Luo, Pengcheng Yang, Wei Wu, Baobao Chang and Zhifang Sui 5985

Style Transformer: Unpaired Text Style Transfer without Disentangled Latent Representation
Ning Dai, Jianze Liang, Xipeng Qiu and Xuanjing Huang 5997

Generating Sentences from Disentangled Syntactic and Semantic Spaces
Yu Bao, Hao Zhou, Shujian Huang, Lei Li, Lili Mou, Olga Vechtomova, Xin-yu Dai and Jiajun Chen 6008

Learning to Control the Fine-grained Sentiment for Story Ending Generation
Fuli Luo, Damai Dai, Pengcheng Yang, Tianyu Liu, Baobao Chang, Zhifang Sui and Xu Sun 6020

Self-Attention Architectures for Answer-Agnostic Neural Question Generation
Thomas Scialom, Benjamin Piwowarski and Jacopo Staiano 6027

Unsupervised Paraphrasing without Translation
Aurko Roy and David Grangier 6033

Storyboarding of Recipes: Grounded Contextual Generation
Khyathi Chandu, Eric Nyberg and Alan W Black 6040

Negative Lexically Constrained Decoding for Paraphrase Generation
Tomoyuki Kajiwara 6047

Large-Scale Transfer Learning for Natural Language Generation
Sergey Golovanov, Rauf Kurbanov, Sergey Nikolenko, Kyryl Truskovskyi, Alexander Tselousov and Thomas Wolf 6053

Automatic Grammatical Error Correction for Sequence-to-sequence Text Generation: An Empirical Study
Tao Ge, Xingxing Zhang, Furu Wei and Ming Zhou 6059

Improving the Robustness of Question Answering Systems to Question Paraphrasing
Wee Chung Gan and Hwee Tou Ng 6065

RankQA: Neural Question Answering with Answer Re-Ranking
Bernhard Kratzwald, Anna Eigenmann and Stefan Feuerriegel 6076

Latent Retrieval for Weakly Supervised Open Domain Question Answering
Kenton Lee, Ming-Wei Chang and Kristina Toutanova 6086

Multi-hop Reading Comprehension through Question Decomposition and Rescoring
Sewon Min, Victor Zhong, Luke Zettlemoyer and Hannaneh Hajishirzi 6097

Combining Knowledge Hunting and Neural Language Models to Solve the Winograd Schema Challenge
Ashok Prakash, Arpit Sharma, Arindam Mitra and Chitta Baral 6110

Careful Selection of Knowledge to Solve Open Book Question Answering
Pratyay Banerjee, Kuntal Kumar Pal, Arindam Mitra and Chitta Baral 6120

Learning Representation Mapping for Relation Detection in Knowledge Base Question Answering
Peng Wu, Shujian Huang, Rongxiang Weng, Zaixiang Zheng, Jianbing Zhang, Xiaohui Yan and Jiajun Chen 6130

Dynamically Fused Graph Network for Multi-hop Reasoning
Lin Qiu, Yunxuan Xiao, Yanru Qu, Hao Zhou, Lei Li, Weinan Zhang and Yong Yu 6140

NLProlog: Reasoning with Weak Unification for Question Answering in Natural Language
Leon Weber, Pasquale Minervini, Jannes Münchmeyer, Ulf Leser and Tim Rocktäschel 6151

Modeling Intra-Relation in Math Word Problems with Different Functional Multi-Head Attentions
Jierui Li, Lei Wang, Jipeng Zhang, Yan Wang, Bing Tian Dai and Dongxiang Zhang 6162

Synthetic QA Corpora Generation with Roundtrip Consistency
Chris Alberti, Daniel Andor, Emily Pitler, Jacob Devlin and Michael Collins 6168

Are Red Roses Red? Evaluating Consistency of Question-Answering Models
Marco Tulio Ribeiro, Carlos Guestrin and Sameer Singh 6174

MC^2: Multi-perspective Convolutional Cube for Conversational Machine Reading Comprehension
Xuanyu Zhang 6185

Reducing Word Omission Errors in Neural Machine Translation: A Contrastive Learning Approach
Zonghan Yang, Yong Cheng, Yang Liu and Maosong Sun 6191

Exploiting Sentential Context for Neural Machine Translation
Xing Wang, Zhaopeng Tu, Longyue Wang and Shuming Shi 6197

Wetin dey with these comments? Modeling Sociolinguistic Factors Affecting Code-switching Behavior in Nigerian Online Discussions
Innocent Ndubuisi-Obi, Sayan Ghosh and David Jurgens 6204

Accelerating Sparse Matrix Operations in Neural Networks on Graphics Processing Units
Arturo Argueta and David Chiang 6215

An Automated Framework for Fast Cognate Detection and Bayesian Phylogenetic Inference in Computational Historical Linguistics
Taraka Rama and Johann-Mattis List 6225

Sentence Centrality Revisited for Unsupervised Summarization
Hao Zheng and Mirella Lapata 6236

Discourse Representation Parsing for Sentences and Documents
Jiangming Liu, Shay B. Cohen and Mirella Lapata 6248

Inducing Document Structure for Aspect-based Summarization
Lea Frermann and Alexandre Klementiev 6263

Incorporating Priors with Feature Attribution on Text Classification
Frederick Liu and Besim Avci 6274

Matching Article Pairs with Graphical Decomposition and Convolutions
Bang Liu, Di Niu, Haojie Wei, Jinghong Lin, Yancheng He, Kunfeng Lai and Yu Xu 6284

Hierarchical Transfer Learning for Multi-label Text Classification
Siddhartha Banerjee, Cem Akkaya, Francisco Perez-Sorrosal and Kostas Tsioutsiouliklis 6295

Bias Analysis and Mitigation in the Evaluation of Authorship Verification
Janek Bevendorff, Matthias Hagen, Benno Stein and Martin Potthast 6301

Numeracy-600K: Learning Numeracy for Detecting Exaggerated Information in Market Comments
Chung-Chi Chen, Hen-Hsen Huang, Hiroya Takamura and Hsin-Hsi Chen 6307

Large-Scale Multi-Label Text Classification on EU Legislation
Ilias Chalkidis, Emmanouil Fergadiotis, Prodromos Malakasiotis and Ion Androutsopoulos . . 6314

Why Didn't You Listen to Me? Comparing User Control of Human-in-the-Loop Topic Models
Varun Kumar, Alison Smith-Renner, Leah Findlater, Kevin Seppi and Jordan Boyd-Graber . . 6323

Encouraging Paragraph Embeddings to Remember Sentence Identity Improves Classification
Tu Vu and Mohit Iyyer 6331

A Multi-Task Architecture on Relevance-based Neural Query Translation
Sheikh Muhammad Sarwar, Hamed Bonab and James Allan 6339

Topic Modeling with Wasserstein Autoencoders
Feng Nan, Ran Ding, Ramesh Nallapati and Bing Xiang 6345

Dense Procedure Captioning in Narrated Instructional Videos
Botian Shi, Lei Ji, Yaobo Liang, Nan Duan, Peng Chen, Zhendong Niu and Ming Zhou 6382

Latent Variable Model for Multi-modal Translation
Iacer Calixto, Miguel Rios and Wilker Aziz 6392

Identifying Visible Actions in Lifestyle Vlogs
Oana Ignat, Laura Burdick, Jia Deng and Rada Mihalcea 6406

A Corpus for Reasoning about Natural Language Grounded in Photographs
Alane Suhr, Stephanie Zhou, Ally Zhang, Iris Zhang, Huajun Bai and Yoav Artzi 6418

Learning to Discover, Ground and Use Words with Segmental Neural Language Models
Kazuya Kawakami, Chris Dyer and Phil Blunsom 6429

What Should I Ask? Using Conversationally Informative Rewards for Goal-oriented Visual Dialog.
Pushkar Shukla, Carlos Elmadjian, Richika Sharan, Vivek Kulkarni, Matthew Turk and William Yang Wang 6442

Symbolic Inductive Bias for Visually Grounded Learning of Spoken Language
Grzegorz Chrupała 6452

Multi-step Reasoning via Recurrent Dual Attention for Visual Dialog
Zhe Gan, Yu Cheng, Ahmed Kholy, Linjie Li, Jingjing Liu and Jianfeng Gao 6463

Lattice Transformer for Speech Translation
Pei Zhang, Niyu Ge, Boxing Chen and Kai Fan 6475

Informative Image Captioning with External Sources of Information
Sanqiang Zhao, Piyush Sharma, Tomer Levinboim and Radu Soricut 6485

CoDraw: Collaborative Drawing as a Testbed for Grounded Goal-driven Communication
Jin-Hwa Kim, Nikita Kitaev, Xinlei Chen, Marcus Rohrbach, Byoung-Tak Zhang, Yuandong Tian, Dhruv Batra and Devi Parikh 6495

Bridging by Word: Image Grounded Vocabulary Construction for Visual Captioning
Zhihao Fan, Zhongyu Wei, Siyuan Wang and Xuanjing Huang 6514

Distilling Translations with Visual Awareness
Julia Ive, Pranava Madhyastha and Lucia Specia 6525

VIFIDEL: Evaluating the Visual Fidelity of Image Descriptions
Pranava Madhyastha, Josiah Wang and Lucia Specia 6539

Are You Looking? Grounding to Multiple Modalities in Vision-and-Language Navigation
Ronghang Hu, Daniel Fried, Anna Rohrbach, Dan Klein, Trevor Darrell and Kate Saenko ... 6551

Multimodal Transformer for Unaligned Multimodal Language Sequences
Yao-Hung Hubert Tsai, Shaojie Bai, Paul Pu Liang, J. Zico Kolter, Louis-Philippe Morency and Ruslan Salakhutdinov 6558

Show, Describe and Conclude: On Exploiting the Structure Information of Chest X-ray Reports
Baoyu Jing, Zeya Wang and Eric Xing 6570

Visual Story Post-Editing
Ting-Yao Hsu, Chieh-Yang Huang, Yen-Chia Hsu and Ting-Hao Huang 6581

Multimodal Abstractive Summarization for How2 Videos
Shruti Palaskar, Jindřich Libovický, Spandana Gella and Florian Metze 6587

Learning to Relate from Captions and Bounding Boxes
Sarthak Garg, Joel Ruben Antony Moniz, Anshu Aviral and Priyatham Bollimpalli 6597