56th Annual Meeting of the Association for Computational Linguistics (ACL 2018)

Long Papers

Melbourne, Australia
15 – 20 July 2018

Volume 1 of 4

ISBN: 978-1-5108-6705-5

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© (2018) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2018)

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

Probabilistic FastText for Multi-Sense Word Embeddings
Ben Athiwaratkun, Andrew Wilson and Anima Anandkumar 1

A La Carte Embedding: Cheap but Effective Induction of Semantic Feature Vectors
Mikhail Khodak, Nikunj Saunshi, Yingyu Liang, Tengyu Ma, Brandon Stewart and Sanjeev Arora 12

Unsupervised Learning of Distributional Relation Vectors
Shoaib Jameel, Zied Bouraoui and Steven Schockaert 23

Explicit Retrofitting of Distributional Word Vectors
Goran Glavaš and Ivan Vulić 34

Unsupervised Neural Machine Translation with Weight Sharing
Zhen Yang, Wei Chen, Feng Wang and Bo Xu 46

Triangular Architecture for Rare Language Translation
Shuo Ren, Wenhu Chen, Shujie Liu, Mu Li, Ming Zhou and Shuai Ma 56

Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates
Taku Kudo 66

The Best of Both Worlds: Combining Recent Advances in Neural Machine Translation
Mia Xu Chen, Orhan Firat, Ankur Bapna, Melvin Johnson, Wolfgang Macherey, George Foster, Llion Jones, Mike Schuster, Noam Shazeer, Niki Parmar, Ashish Vaswani, Jakob Uszkoreit, Lukasz Kaiser, Zhifeng Chen, Yonghui Wu and Macduff Hughes 76

Ultra-Fine Entity Typing
Eunsol Choi, Omer Levy, Yejin Choi and Luke Zettlemoyer 87

Hierarchical Losses and New Resources for Fine-grained Entity Typing and Linking
Shikhar Murty, Patrick Verga, Luke Vilnis, Irena Radovanovic and Andrew McCallum 97

Improving Knowledge Graph Embedding Using Simple Constraints
Boyang Ding, Quan Wang, Bin Wang and Li Guo 110

Towards Understanding the Geometry of Knowledge Graph Embeddings
Chandrahas -, Aditya Sharma and Partha Talukdar 122

A Unified Model for Extractive and Abstractive Summarization using Inconsistency Loss
Wan-Ting Hsu, Chieh-Kai Lin, Ming-Ying Lee, Kerui Min, Jing Tang and Min Sun 132

Extractive Summarization with SWAP-NET: Sentences and Words from Alternating Pointer Networks
Aishwarya Jadhav and Vaibhav Rajan 142

Retrieve, Rerank and Rewrite: Soft Template Based Neural Summarization
Ziqiang Cao, Wenjie Li, Sujian Li and Furu Wei 152

Simple and Effective Text Simplification Using Semantic and Neural Methods
Elior Sulem, Omri Abend and Ari Rappoport 162

Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words
Saif Mohammad 174

Comprehensive Supersense Disambiguation of English Prepositions and Possessives
Nathan Schneider, Jena D. Hwang, Vivek Srikumar, Jakob Prange, Austin Blodgett, Sarah R. Moeller, Aviram Stern, Adi Bitan and Omri Abend 185

A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature
Benjamin Nye, Junyi Jessy Li, Roma Patel, Yinfei Yang, Iain Marshall, Ani Nenkova and Byron Wallace 197

Efficient Online Scalar Annotation with Bounded Support
Keisuke Sakaguchi and Benjamin Van Durme 208

Neural Argument Generation Augmented with Externally Retrieved Evidence
Xinyu Hua and Lu Wang 219

A Stylometric Inquiry into Hyperpartisan and Fake News
Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff and Benno Stein 231

Retrieval of the Best Counterargument without Prior Topic Knowledge
Henning Wachsmuth, Shahbaz Syed and Benno Stein 241

LinkNBed: Multi-Graph Representation Learning with Entity Linkage
Rakshit Trivedi, Bunyamin Sisman, Xin Luna Dong, Christos Faloutsos, Jun Ma and Hongyuan Zha 252

Probabilistic Embedding of Knowledge Graphs with Box Lattice Measures
Luke Vilnis, Xiang Li, Shikhar Murty and Andrew McCallum 263

Graph-to-Sequence Learning using Gated Graph Neural Networks
Daniel Beck, Gholamreza Haffari and Trevor Cohn 273

Sharp Nearby, Fuzzy Far Away: How Neural Language Models Use Context
Urvashi Khandelwal, He He, Peng Qi and Dan Jurafsky 284

Bridging CNNs, RNNs, and Weighted Finite-State Machines
Roy Schwartz, Sam Thomson and Noah A. Smith 295

Zero-shot Learning of Classifiers from Natural Language Quantification
Shashank Srivastava, Igor Labutov and Tom Mitchell 306

Sentence-State LSTM for Text Representation
Yue Zhang, Qi Liu and Linfeng Song 317

Universal Language Model Fine-tuning for Text Classification
Jeremy Howard and Sebastian Ruder 328

Evaluating neural network explanation methods using hybrid documents and morphosyntactic agreement
Nina Poerner, Hinrich Schütze and Benjamin Roth 340

Improving Text-to-SQL Evaluation Methodology
Catherine Finegan-Dollak, Jonathan K. Kummerfeld, Li Zhang, Karthik Ramanathan, Sesh Sadasivam, Rui Zhang and Dragomir Radev 351

Semantic Parsing with Syntax- and Table-Aware SQL Generation
Yibo Sun, Duyu Tang, Nan Duan, Jianshu Ji, Guihong Cao, Xiaocheng Feng, Bing Qin, Ting Liu and Ming Zhou 361

Multitask Parsing Across Semantic Representations
Daniel Hershcovich, Omri Abend and Ari Rappoport 373

Character-Level Models versus Morphology in Semantic Role Labeling
Gozde Gul Sahin and Mark Steedman 386

AMR Parsing as Graph Prediction with Latent Alignment
Chunchuan Lyu and Ivan Titov 397

Accurate SHRG-Based Semantic Parsing
Yufei Chen, Weiwei Sun and Xiaojun Wan 408

Using Intermediate Representations to Solve Math Word Problems
Danqing Huang, Jin-Ge Yao, Chin-Yew Lin, Qingyu Zhou and Jian Yin 419

Discourse Representation Structure Parsing
Jiangming Liu, Shay B. Cohen and Mirella Lapata 429

Baseline Needs More Love: On Simple Word-Embedding-Based Models and Associated Pooling Mechanisms
Dinghan Shen, Guoyin Wang, Wenlin Wang, Martin Renqiang Min, Qinliang Su, Yizhe Zhang, Chunyuan Li, Ricardo Henao and Lawrence Carin 440

ParaNMT-50M: Pushing the Limits of Paraphrastic Sentence Embeddings with Millions of Machine Translations
John Wieting and Kevin Gimpel 451

Event2Mind: Commonsense Inference on Events, Intents, and Reactions
Hannah Rashkin, Maarten Sap, Emily Allaway, Noah A. Smith and Yejin Choi 463

Neural Adversarial Training for Semi-supervised Japanese Predicate-argument Structure Analysis
Shuhei Kurita, Daisuke Kawahara and Sadao Kurohashi 474

Improving Event Coreference Resolution by Modeling Correlations between Event Coreference Chains and Document Topic Structures
Prafulla Kumar Choubey and Ruihong Huang 485

DSGAN: Generative Adversarial Training for Distant Supervision Relation Extraction
Pengda Qin, Weiran XU and William Yang Wang 496

Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism
Xiangrong Zeng, Daojian Zeng, Shizhu He, Kang Liu and Jun Zhao 506

Self-regulation: Employing a Generative Adversarial Network to Improve Event Detection
Yu Hong, Wenxuan Zhou, jingli zhang, Guodong Zhou and Qiaoming Zhu 515

Context-Aware Neural Model for Temporal Information Extraction
Yuanliang Meng and Anna Rumshisky 527

Temporal Event Knowledge Acquisition via Identifying Narratives
Wenlin Yao and Ruihong Huang 537

Textual Deconvolution Saliency (TDS) : a deep tool box for linguistic analysis
Laurent Vanni, Mélanie Ducoffe, Carlos Aguilar, Frederic Precioso and Damon Mayaffre 548

Coherence Modeling of Asynchronous Conversations: A Neural Entity Grid Approach
Shafiq Joty, Muhammad Tasnim Mohiuddin and Dat Tien Nguyen 558

Deep Reinforcement Learning for Chinese Zero Pronoun Resolution
Qingyu Yin, Yu Zhang, Wei-Nan Zhang, Ting Liu and William Yang Wang 569

Entity-Centric Joint Modeling of Japanese Coreference Resolution and Predicate Argument Structure Analysis
Tomohide Shibata and Sadao Kurohashi 579

Constraining MGbank: Agreement, L-Selection and Supertagging in Minimalist Grammars
John Torr 590

Not that much power: Linguistic alignment is influenced more by low-level linguistic features rather than social power
Yang Xu, Jeremy Cole and David Reitter 601

TutorialBank: A Manually-Collected Corpus for Prerequisite Chains, Survey Extraction and Resource Recommendation
Alexander Fabbri, Irene Li, Prawat Trairatvorakul, Yijiao He, Weitai Ting, Robert Tung, Caitlin Westerfield and Dragomir Radev 611

Give Me More Feedback: Annotating Argument Persuasiveness and Related Attributes in Student Essays
Winston Carlile, Nishant Gurrapadi, Zixuan Ke and Vincent Ng 621

Inherent Biases in Reference-based Evaluation for Grammatical Error Correction
Leshem Choshen and Omri Abend 632

The price of debiasing automatic metrics in natural language evalaution
Arun Chaganty, Stephen Mussmann and Percy Liang 643

Neural Document Summarization by Jointly Learning to Score and Select Sentences
Qingyu Zhou, Nan Yang, Furu Wei, Shaohan Huang, Ming Zhou and Tiejun Zhao 654

Unsupervised Abstractive Meeting Summarization with Multi-Sentence Compression and Budgeted Submodular Maximization
Guokan Shang, Wensi Ding, Zekun Zhang, Antoine Tixier, Polykarpos Meladianos, Michalis Vazirgiannis and Jean-Pierre Lorré 664

Fast Abstractive Summarization with Reinforce-Selected Sentence Rewriting
Yen-Chun Chen and Mohit Bansal 675

Soft Layer-Specific Multi-Task Summarization with Entailment and Question Generation
Han Guo, Ramakanth Pasunuru and Mohit Bansal 687

Modeling and Prediction of Online Product Review Helpfulness: A Survey
Gerardo Ocampo Diaz and Vincent Ng 698

Mining Cross-Cultural Differences and Similarities in Social Media
Bill Yuchen Lin, Frank F. Xu, Kenny Zhu and Seung-won Hwang 709

Classification of Moral Foundations in Microblog Political Discourse
Kristen Johnson and Dan Goldwasser 720

Coarse-to-Fine Decoding for Neural Semantic Parsing
Li Dong and Mirella Lapata 731

Confidence Modeling for Neural Semantic Parsing
Li Dong, Chris Quirk and Mirella Lapata 743

StructVAE: Tree-structured Latent Variable Models for Semi-supervised Semantic Parsing
Pengcheng Yin, Chunting Zhou, Junxian He and Graham Neubig 754

Sequence-to-Action: End-to-End Semantic Graph Generation for Semantic Parsing
Bo Chen, Le Sun and Xianpei Han 766

On the Limitations of Unsupervised Bilingual Dictionary Induction
Anders Søgaard, Sebastian Ruder and Ivan Vulić 778

A robust self-learning method for fully unsupervised cross-lingual mappings of word embeddings
Mikel Artetxe, Gorka Labaka and Eneko Agirre 789

A Multi-lingual Multi-task Architecture for Low-resource Sequence Labeling
Ying Lin, Shengqi Yang, Veselin Stoyanov and Heng Ji 799

Two Methods for Domain Adaptation of Bilingual Tasks: Delightfully Simple and Broadly Applicable
Viktor Hangya, Fabienne Braune, Alexander Fraser and Hinrich Schütze 810

Knowledgeable Reader: Enhancing Cloze-Style Reading Comprehension with External Commonsense Knowledge
Todor Mihaylov and Anette Frank 821

Multi-Relational Question Answering from Narratives: Machine Reading and Reasoning in Simulated Worlds
Igor Labutov, Bishan Yang, Anusha Prakash and Amos Azaria 833

Simple and Effective Multi-Paragraph Reading Comprehension
Christopher Clark and Matt Gardner 845

Semantically Equivalent Adversarial Rules for Debugging NLP models
Marco Tulio Ribeiro, Sameer Singh and Carlos Guestrin 856

Style Transfer Through Back-Translation
Shrimai Prabhumoye, Yulia Tsvetkov, Ruslan Salakhutdinov and Alan W Black 866

Generating Fine-Grained Open Vocabulary Entity Type Descriptions
Rajarshi Bhowmik and Gerard de Melo 877

Hierarchical Neural Story Generation
Angela Fan, Mike Lewis and Yann Dauphin 889

No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling
Xin Wang, Wenhu Chen, Yuan-Fang Wang and William Yang Wang 899

Bridging Languages through Images with Deep Partial Canonical Correlation Analysis
Guy Rotman, Ivan Vulić and Roi Reichart 910

Illustrative Language Understanding: Large-Scale Visual Grounding with Image Search
Jamie Kiros, William Chan and Geoffrey Hinton 922

What Action Causes This? Towards Naive Physical Action-Effect Prediction
Qiaozi Gao, Shaohua Yang, Joyce Chai and Lucy Vanderwende 934

Transformation Networks for Target-Oriented Sentiment Classification
Xin Li, Lidong Bing, Wai Lam and Bei Shi 946

Target-Sensitive Memory Networks for Aspect Sentiment Classification
Shuai Wang, Sahisnu Mazumder, Bing Liu, Mianwei Zhou and Yi Chang 957

Identifying Transferable Information Across Domains for Cross-domain Sentiment Classification
Raksha Sharma, Pushpak Bhattacharyya, Sandipan Dandapat and Himanshu Sharad Bhatt 968

Unpaired Sentiment-to-Sentiment Translation: A Cycled Reinforcement Learning Approach
Jingjing Xu, Xu SUN, Qi Zeng, Xiaodong Zhang, Xuancheng Ren, Houfeng Wang and Wenjie Li 979

Discourse Marker Augmented Network with Reinforcement Learning for Natural Language Inference
Boyuan Pan, Yazheng Yang, Zhou Zhao, Yueting Zhuang, Deng Cai and Xiaofei He 989

Working Memory Networks: Augmenting Memory Networks with a Relational Reasoning Module
Juan Pavez, Hector Alllende and Hector Allende-Cid 1000

Reasoning with Sarcasm by Reading In-Between
Yi Tay, Anh Tuan Luu, Siu Cheung Hui and Jian Su 1010

Adversarial Contrastive Estimation
Avishek Joey Bose, huan ling and Yanshuai Cao 1021

Adaptive Scaling for Sparse Detection in Information Extraction
Hongyu Lin, Yaojie Lu, Xianpei Han and Le Sun 1033

Strong Baselines for Neural Semi-Supervised Learning under Domain Shift
Sebastian Ruder and Barbara Plank 1044

Fluency Boost Learning and Inference for Neural Grammatical Error Correction
Tao Ge, Furu Wei and Ming Zhou 1055

A Neural Architecture for Automated ICD Coding
Pengtao Xie and Eric Xing 1066

Domain Adaptation with Adversarial Training and Graph Embeddings
Firoj Alam, Shafiq Joty and Muhammad Imran 1077

TDNN: A Two-stage Deep Neural Network for Prompt-independent Automated Essay Scoring
Cancan Jin, Ben He, Kai Hui and Le Sun 1088

Unsupervised Discrete Sentence Representation Learning for Interpretable Neural Dialog Generation
Tiancheng Zhao, Kyusong Lee and Maxine Eskenazi 1098

Learning to Control the Specificity in Neural Response Generation
Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, Jun Xu and Xueqi Cheng 1108

Multi-Turn Response Selection for Chatbots with Deep Attention Matching Network
Xiangyang Zhou, Lu Li, Daxiang Dong, Yi Liu, Ying Chen, Wayne Xin Zhao, Dianhai Yu and Hua Wu 1118

MojiTalk: Generating Emotional Responses at Scale
Xianda Zhou and William Yang Wang 1128

Taylor's law for Human Linguistic Sequences
Tatsuru Kobayashi and Kumiko Tanaka-Ishii 1138

A Framework for Representing Language Acquisition in a Population Setting
Jordan Kodner and Christopher Cerezo Falco 1149

Prefix Lexicalization of Synchronous CFGs using Synchronous TAG
Logan Born and Anoop Sarkar 1160

Straight to the Tree: Constituency Parsing with Neural Syntactic Distance
Yikang Shen, Zhouhan Lin, Athul Paul Jacob, Alessandro Sordoni, Aaron Courville and Yoshua Bengio 1171

Gaussian Mixture Latent Vector Grammars
Yanpeng Zhao, Liwen Zhang and Kewei Tu 1181

Extending a Parser to Distant Domains Using a Few Dozen Partially Annotated Examples
Vidur Joshi, Matthew Peters and Mark Hopkins 1190

Paraphrase to Explicate: Revealing Implicit Noun-Compound Relations
Vered Shwartz and Ido Dagan 1200

Searching for the X-Factor: Exploring Corpus Subjectivity for Word Embeddings
Maksim Tkachenko, Chong Cher Chia and Hady Lauw 1212

Word Embedding and WordNet Based Metaphor Identification and Interpretation
Rui Mao, Chenghua Lin and Frank Guerin 1222

Incorporating Latent Meanings of Morphological Compositions to Enhance Word Embeddings
Yang Xu, Jiawei Liu, Wei Yang and Liusheng Huang 1232

A Stochastic Decoder for Neural Machine Translation
Philip Schulz, Wilker Aziz and Trevor Cohn 1243

Forest-Based Neural Machine Translation
Chunpeng Ma, Akihiro Tamura, Masao Utiyama, Tiejun Zhao and Eiichiro Sumita 1253

Context-Aware Neural Machine Translation Learns Anaphora Resolution
Elena Voita, Pavel Serdyukov, Rico Sennrich and Ivan Titov 1264

Document Context Neural Machine Translation with Memory Networks
Sameen Maruf and Gholamreza Haffari 1275

Which Melbourne? Augmenting Geocoding with Maps
Milan Gritta, Mohammad Taher Pilehvar and Nigel Collier 1285

Learning Prototypical Goal Activities for Locations
Tianyu Jiang and Ellen Riloff 1297

Guess Me if You Can: Acronym Disambiguation for Enterprises
Yang Li, Bo Zhao, Ariel Fuxman and Fangbo Tao 1308

A Multi-Axis Annotation Scheme for Event Temporal Relations
Qiang Ning, Hao Wu and Dan Roth 1318

Exemplar Encoder-Decoder for Neural Conversation Generation
Gaurav Pandey, Danish Contractor, Vineet Kumar and Sachindra Joshi 1329

DialSQL: Dialogue Based Structured Query Generation
Izzeddin Gur, Semih Yavuz, Yu Su and Xifeng Yan 1339

Conversations Gone Awry: Detecting Early Signs of Conversational Failure
Justine Zhang, Jonathan Chang, Cristian Danescu-Niculescu-Mizil, Lucas Dixon, Yiqing Hua, Dario Taraborelli and Nithum Thain 1350

Are BLEU and Meaning Representation in Opposition?
Ondřej Cífka and Ondřej Bojar 1362

Automatic Metric Validation for Grammatical Error Correction
Leshem Choshen and Omri Abend 1372

The Hitchhiker's Guide to Testing Statistical Significance in Natural Language Processing
Rotem Dror, Gili Baumer, Segev Shlomov and Roi Reichart 1383

Distilling Knowledge for Search-based Structured Prediction
Yijia Liu, Wanxiang Che, Huaipeng Zhao, Bing Qin and Ting Liu 1393

Stack-Pointer Networks for Dependency Parsing
Xuezhe Ma, Zecong Hu, Jingzhou Liu, Nanyun Peng, Graham Neubig and Eduard Hovy 1403

Twitter Universal Dependency Parsing for African-American and Mainstream American English
Su Lin Blodgett, Johnny Wei and Brendan O'Connor 1415

LSTMs Can Learn Syntax-Sensitive Dependencies Well, But Modeling Structure Makes Them Better
Adhiguna Kuncoro, Chris Dyer, John Hale, Dani Yogatama, Stephen Clark and Phil Blunsom 1426

Sequicity: Simplifying Task-oriented Dialogue Systems with Single Sequence-to-Sequence Architectures
Wenqiang Lei, Xisen Jin, Min-Yen Kan, Zhaochun Ren, Xiangnan He and Dawei Yin 1437

An End-to-end Approach for Handling Unknown Slot Values in Dialogue State Tracking
Puyang Xu and Qi Hu 1448

Global-Locally Self-Attentive Encoder for Dialogue State Tracking
Victor Zhong, Caiming Xiong and Richard Socher 1458

Mem2Seq: Effectively Incorporating Knowledge Bases into End-to-End Task-Oriented Dialog Systems
Andrea Madotto, Chien-Sheng Wu and Pascale Fung 1468

Tailored Sequence to Sequence Models to Different Conversation Scenarios
Hainan Zhang, Yanyan Lan, Jiafeng Guo, Jun Xu and Xueqi Cheng 1479

Knowledge Diffusion for Neural Dialogue Generation
Shuman Liu, Hongshen Chen, Zhaochun Ren, Yang Feng, Qun Liu and Dawei Yin 1489

Generating Informative Responses with Controlled Sentence Function
Pei Ke, Jian Guan, Minlie Huang and xiaoyan zhu 1499

Sentiment Adaptive End-to-End Dialog Systems
Weiyan Shi and Zhou Yu 1509

Embedding Learning Through Multilingual Concept Induction
Philipp Dufter, Mengjie Zhao, Martin Schmitt, Alexander Fraser and Hinrich Schütze 1520

Isomorphic Transfer of Syntactic Structures in Cross-Lingual NLP
Edoardo Maria Ponti, Roi Reichart, Anna Korhonen and Ivan Vulić 1531

Language Modeling for Code-Mixing: The Role of Linguistic Theory based Synthetic Data
Adithya Pratapa, Gayatri Bhat, Monojit Choudhury, Sunayana Sitaram, Sandipan Dandapat and Kalika Bali 1543

Chinese NER Using Lattice LSTM
Yue Zhang and Jie Yang 1554

Nugget Proposal Networks for Chinese Event Detection
Hongyu Lin, Yaojie Lu, Xianpei Han and Le Sun 1565

Higher-order Relation Schema Induction using Tensor Factorization with Back-off and Aggregation
Madhav Nimishakavi, Manish Gupta and Partha Talukdar 1575

Discovering Implicit Knowledge with Unary Relations
Michael Glass and Alfio Gliozzo 1585

Improving Entity Linking by Modeling Latent Relations between Mentions
Phong Le and Ivan Titov 1595

Dating Documents using Graph Convolution Networks
Shikhar Vashishth, Shib Sankar Dasgupta, Swayambhu Nath Ray and Partha Talukdar 1605

A Graph-to-Sequence Model for AMR-to-Text Generation
Linfeng Song, Yue Zhang, Zhiguo Wang and Daniel Gildea 1616

GTR-LSTM: A Triple Encoder for Sentence Generation from RDF Data
Bayu Distiawan Trisedya, Jianzhong Qi, Rui Zhang and Wei Wang 1627

Learning to Write with Cooperative Discriminators
Ari Holtzman, Jan Buys, Maxwell Forbes, Antoine Bosselut, David Golub and Yejin Choi . . . 1638

A Neural Approach to Pun Generation
Zhiwei Yu, Jiwei Tan and Xiaojun Wan 1650

Learning to Generate Move-by-Move Commentary for Chess Games from Large-Scale Social Forum Data
Harsh Jhamtani, Varun Gangal, Eduard Hovy, Graham Neubig and Taylor Berg-Kirkpatrick . . 1661

From Credit Assignment to Entropy Regularization: Two New Algorithms for Neural Sequence Prediction
Zihang Dai, Qizhe Xie and Eduard Hovy 1672

DuoRC: Towards Complex Language Understanding with Paraphrased Reading Comprehension
Amrita Saha, Rahul Aralikatte, Mitesh M. Khapra and Karthik Sankaranarayanan 1683

Stochastic Answer Networks for Machine Reading Comprehension
Xiaodong Liu, Yelong Shen, Kevin Duh and Jianfeng Gao 1694

Multi-Granularity Hierarchical Attention Fusion Networks for Reading Comprehension and Question Answering
Wei Wang, Ming Yan and Chen Wu 1705

Joint Training of Candidate Extraction and Answer Selection for Reading Comprehension
Zhen Wang, Jiachen Liu, Xinyan Xiao, Yajuan Lyu and Tian Wu 1715

Efficient and Robust Question Answering from Minimal Context over Documents
Sewon Min, Victor Zhong, Richard Socher and Caiming Xiong 1725

Denoising Distantly Supervised Open-Domain Question Answering
Yankai Lin, Haozhe Ji, Zhiyuan Liu and Maosong Sun 1736

Question Condensing Networks for Answer Selection in Community Question Answering
Wei Wu, Xu SUN and Houfeng WANG 1746

Towards Robust Neural Machine Translation
Yong Cheng, Zhaopeng Tu, Fandong Meng, Junjie Zhai and Yang Liu 1756

Attention Focusing for Neural Machine Translation by Bridging Source and Target Embeddings
Shaohui Kuang, Junhui Li, António Branco, Weihua Luo and Deyi Xiong 1767

Reliability and Learnability of Human Bandit Feedback for Sequence-to-Sequence Reinforcement Learning
Julia Kreutzer, Joshua Uyheng and Stefan Riezler 1777

Accelerating Neural Transformer via an Average Attention Network
Biao Zhang, Deyi Xiong and jinsong su 1789

How Much Attention Do You Need? A Granular Analysis of Neural Machine Translation Architectures
Tobias Domhan 1799

Weakly Supervised Semantic Parsing with Abstract Examples
Omer Goldman, Veronica Latcinnik, Ehud Nave, Amir Globerson and Jonathan Berant 1809

Improving a Neural Semantic Parser by Counterfactual Learning from Human Bandit Feedback
Carolin Lawrence and Stefan Riezler 1820

AMR dependency parsing with a typed semantic algebra
Jonas Groschwitz, Matthias Lindemann, Meaghan Fowlie, Mark Johnson and Alexander Koller 1831

Sequence-to-sequence Models for Cache Transition Systems
Xiaochang Peng, Linfeng Song, Daniel Gildea and Giorgio Satta 1842

Batch IS NOT Heavy: Learning Word Representations From All Samples
Xin Xin, Fajie Yuan, Xiangnan He and Joemon M Jose 1853

Backpropagating through Structured Argmax using a SPIGOT
Hao Peng, Sam Thomson and Noah A. Smith 1863

Learning How to Actively Learn: A Deep Imitation Learning Approach
Ming Liu, Wray Buntine and Gholamreza Haffari 1874

Training Classifiers with Natural Language Explanations
Braden Hancock, Paroma Varma, Stephanie Wang, Martin Bringmann, Percy Liang and Christopher Ré 1884

Did the Model Understand the Question?
Pramod Kaushik Mudrakarta, Ankur Taly, Mukund Sundararajan and Kedar Dhamdhere 1896

Harvesting Paragraph-level Question-Answer Pairs from Wikipedia
Xinya Du and Claire Cardie 1907

Multi-Passage Machine Reading Comprehension with Cross-Passage Answer Verification
Yizhong Wang, Kai Liu, Jing Liu, Wei He, Yajuan Lyu, Hua Wu, Sujian Li and Haifeng Wang 1918

Language Generation via DAG Transduction
Yajie Ye, Weiwei Sun and Xiaojun Wan 1928

A Distributional and Orthographic Aggregation Model for English Derivational Morphology
Daniel Deutsch, John Hewitt and Dan Roth 1938

Deep-speare: A joint neural model of poetic language, meter and rhyme
Jey Han Lau, Trevor Cohn, Timothy Baldwin, Julian Brooke and Adam Hammond 1948

NeuralREG: An end-to-end approach to referring expression generation
Thiago Castro Ferreira, Diego Moussallem, Ákos Kádár, Sander Wubben and Emiel Krahmer 1959

Stock Movement Prediction from Tweets and Historical Prices
Yumo Xu and Shay B. Cohen 1970

Rumor Detection on Twitter with Tree-structured Recursive Neural Networks
Jing Ma, Wei Gao and Kam-Fai Wong 1980

Visual Attention Model for Name Tagging in Multimodal Social Media
Di Lu, Leonardo Neves, Vitor Carvalho, Ning Zhang and Heng Ji 1990

Multimodal Named Entity Disambiguation for Noisy Social Media Posts
Seungwhan Moon, Leonardo Neves and Vitor Carvalho 2000

Semi-supervised User Geolocation via Graph Convolutional Networks
Afshin Rahimi, Trevor Cohn and Timothy Baldwin 2009

Document Modeling with External Attention for Sentence Extraction
Shashi Narayan, Ronald Cardenas, Nikos Papasarantopoulos, Shay B. Cohen, Mirella Lapata, Jiangsheng Yu and Yi Chang 2020

Neural Models for Documents with Metadata
Dallas Card, Chenhao Tan and Noah A. Smith 2031

NASH: Toward End-to-End Neural Architecture for Generative Semantic Hashing
Dinghan Shen, Qinliang Su, Paidamoyo Chapfuwa, Wenlin Wang, Guoyin Wang, Ricardo Henao and Lawrence Carin 2041

Large-Scale QA-SRL Parsing
Nicholas FitzGerald, Julian Michael, Luheng He and Luke Zettlemoyer 2051

Syntax for Semantic Role Labeling, To Be, Or Not To Be
Shexia He, Zuchao Li, Hai Zhao and Hongxiao Bai 2061

Situated Mapping of Sequential Instructions to Actions with Single-step Reward Observation
Alane Suhr and Yoav Artzi 2072

Marrying Up Regular Expressions with Neural Networks: A Case Study for Spoken Language Understanding
Bingfeng Luo, Yansong Feng, Zheng Wang, Songfang Huang, Rui Yan and Dongyan Zhao . . 2083

Token-level and sequence-level loss smoothing for RNN language models
Maha ELBAYAD, Laurent Besacier and Jakob Verbeek 2094

Numeracy for Language Models: Evaluating and Improving their Ability to Predict Numbers
Georgios Spithourakis and Sebastian Riedel 2104

To Attend or not to Attend: A Case Study on Syntactic Structures for Semantic Relatedness
Amulya Gupta and Zhu Zhang 2116

What you can cram into a single $&!# vector: Probing sentence embeddings for linguistic properties*
Alexis Conneau, Germán Kruszewski, Guillaume Lample, Loïc Barrault and Marco Baroni . . 2126

Robust Distant Supervision Relation Extraction via Deep Reinforcement Learning
Pengda Qin, Weiran XU and William Yang Wang 2137

Interpretable and Compositional Relation Learning by Joint Training with an Autoencoder
Ryo Takahashi, Ran Tian and Kentaro Inui 2148

Zero-Shot Transfer Learning for Event Extraction
Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel and Clare Voss 2160

Recursive Neural Structural Correspondence Network for Cross-domain Aspect and Opinion Co-Extraction
Wenya Wang and Sinno Jialin Pan 2171

Deep Dyna-Q: Integrating Planning for Task-Completion Dialogue Policy Learning
Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjing Liu and Kam-Fai Wong 2182

Learning to Ask Questions in Open-domain Conversational Systems with Typed Decoders
Yansen Wang, Chenyi Liu, Minlie Huang and Liqiang Nie 2193

Personalizing Dialogue Agents: I have a dog, do you have pets too?
Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela and Jason Weston 2204

Efficient Large-Scale Neural Domain Classification with Personalized Attention
Young-Bum Kim, Dongchan Kim, Anjishnu Kumar and Ruhi Sarikaya 2214

Multimodal Affective Analysis Using Hierarchical Attention Strategy with Word-Level Alignment
Yue Gu, Kangning Yang, Shiyu Fu, Shuhong Chen, Xinyu Li and Ivan Marsic 2225

Multimodal Language Analysis in the Wild: CMU-MOSEI Dataset and Interpretable Dynamic Fusion Graph
AmirAli Bagher Zadeh, Paul Pu Liang, Soujanya Poria, Erik Cambria and Louis-Philippe Morency 2236

Efficient Low-rank Multimodal Fusion With Modality-Specific Factors
Zhun Liu, Ying Shen, Varun Bharadhwaj Lakshminarasimhan, Paul Pu Liang, AmirAli Bagher Zadeh and Louis-Philippe Morency 2247

Discourse Coherence: Concurrent Explicit and Implicit Relations
Hannah Rohde, Alexander Johnson, Nathan Schneider and Bonnie Webber 2257

A Spatial Model for Extracting and Visualizing Latent Discourse Structure in Text
Shashank Srivastava and Nebojsa Jojic 2268

Joint Reasoning for Temporal and Causal Relations
Qiang Ning, Zhili Feng, Hao Wu and Dan Roth 2278

Modeling Naive Psychology of Characters in Simple Commonsense Stories
Hannah Rashkin, Antoine Bosselut, Maarten Sap, Kevin Knight and Yejin Choi 2289

A Deep Relevance Model for Zero-Shot Document Filtering
Chenliang Li, Wei Zhou, Feng Ji, Yu Duan and Haiqing Chen 2300

Disconnected Recurrent Neural Networks for Text Categorization
Baoxin Wang 2311

Joint Embedding of Words and Labels for Text Classification
Guoyin Wang, Chunyuan Li, Wenlin Wang, Yizhe Zhang, Dinghan Shen, Xinyuan Zhang, Ricardo Henao and Lawrence Carin 2321

Neural Sparse Topical Coding
Min Peng, Qianqian Xie, Yanchun Zhang, Hua Wang, Xiuzhen Zhang, Jimin Huang and Gang Tian 2332

Document Similarity for Texts of Varying Lengths via Hidden Topics
Hongyu Gong, Tarek Sakakini, Suma Bhat and JinJun Xiong 2341

Eyes are the Windows to the Soul: Predicting the Rating of Text Quality Using Gaze Behaviour
Sandeep Mathias, Diptesh Kanojia, Kevin Patel, Samarth Agrawal, Abhijit Mishra and Pushpak Bhattacharyya 2352

Multi-Input Attention for Unsupervised OCR Correction
Rui Dong and David Smith 2363

Building Language Models for Text with Named Entities
Md Rizwan Parvez, Saikat Chakraborty, Baishakhi Ray and Kai-Wei Chang 2373

hyperdoc2vec: Distributed Representations of Hypertext Documents
Jialong Han, Yan Song, Wayne Xin Zhao, Shuming Shi and Haisong Zhang 2384

Entity-Duet Neural Ranking: Understanding the Role of Knowledge Graph Semantics in Neural Information Retrieval
Zhenghao Liu, Chenyan Xiong, Maosong Sun and Zhiyuan Liu 2395

Neural Natural Language Inference Models Enhanced with External Knowledge
Qian Chen, Xiaodan Zhu, Zhen-Hua Ling, Diana Inkpen and Si Wei 2406

AdvEntuRe: Adversarial Training for Textual Entailment with Knowledge-Guided Examples
Dongyeop Kang, Tushar Khot, Ashish Sabharwal and Eduard Hovy 2418

Subword-level Word Vector Representations for Korean
Sungjoon Park, Jeongmin Byun, Sion Baek, Yongseok Cho and Alice Oh 2429

Incorporating Chinese Characters of Words for Lexical Sememe Prediction
Huiming Jin, Hao Zhu, Zhiyuan Liu, Ruobing Xie, Maosong Sun, Fen Lin and Leyu Lin 2439

SemAxis: A Lightweight Framework to Characterize Domain-Specific Word Semantics Beyond Sentiment
Jisun An, Haewoon Kwak and Yong-Yeol Ahn 2450

End-to-End Reinforcement Learning for Automatic Taxonomy Induction
Yuning Mao, Xiang Ren, Jiaming Shen, Xiaotao Gu and Jiawei Han 2462

Incorporating Glosses into Neural Word Sense Disambiguation
Fuli Luo, Tianyu Liu, Qiaolin Xia, Baobao Chang and Zhifang Sui 2473

Bilingual Sentiment Embeddings: Joint Projection of Sentiment Across Languages
Jeremy Barnes, Roman Klinger and Sabine Schulte im Walde 2483

Learning Domain-Sensitive and Sentiment-Aware Word Embeddings
Bei Shi, Zihao Fu, Lidong Bing and Wai Lam 2494

Cross-Domain Sentiment Classification with Target Domain Specific Information
Minlong Peng, Qi Zhang, Yu-gang Jiang and Xuanjing Huang 2505

Aspect Based Sentiment Analysis with Gated Convolutional Networks
Wei Xue and Tao Li 2514

A Helping Hand: Transfer Learning for Deep Sentiment Analysis
Xin Dong and Gerard de Melo 2524

Cold-Start Aware User and Product Attention for Sentiment Classification
Reinald Kim Amplayo, Jihyeok Kim, Sua Sung and Seung-won Hwang 2535

Modeling Deliberative Argumentation Strategies on Wikipedia
Khalid Al Khatib, Henning Wachsmuth, Kevin Lang, Jakob Herpel, Matthias Hagen and Benno Stein 2545

Conceptual Captions: A Cleaned, Hypernymed, Image Alt-text Dataset For Automatic Image Captioning
Piyush Sharma, Nan Ding, Sebastian Goodman and Radu Soricut 2556

Learning Translations via Images with a Massively Multilingual Image Dataset
John Hewitt, Daphne Ippolito, Brendan Callahan, Reno Kriz, Derry Tanti Wijaya and Chris Callison-Burch 2566

On the Automatic Generation of Medical Imaging Reports
Baoyu Jing, Pengtao Xie and Eric Xing 2577

Attacking Visual Language Grounding with Adversarial Examples: A Case Study on Neural Image Captioning
Hongge Chen, Huan Zhang, Pin-Yu Chen, Jinfeng Yi and Cho-Jui Hsieh 2587

Think Visually: Question Answering through Virtual Imagery
Ankit Goyal, Jian Wang and Jia Deng 2598

Interactive Language Acquisition with One-shot Visual Concept Learning through a Conversational Game
Haichao Zhang, Haonan Yu and Wei Xu 2609

A Purely End-to-End System for Multi-speaker Speech Recognition
Hiroshi Seki, Takaaki Hori, Shinji Watanabe, Jonathan Le Roux and John R Hershey 2620

The OSU Realizer for SRST '18: Neural Sequence-to-Sequence Inflection
David King, Michael White 2631

Morphosyntactic Tagging with a Meta-BiLSTM Model over Context Sensitive Token Encodings
Bernd Bohnet, Ryan McDonald, Gonçalo Simões, Daniel Andor, Emily Pitler and Joshua Maynez 2641

Neural Factor Graph Models for Cross-lingual Morphological Tagging
Chaitanya Malaviya, Matthew R. Gormley and Graham Neubig 2652

Global Transition-based Non-projective Dependency Parsing
Carlos Gómez-Rodríguez, Tianze Shi and Lillian Lee 2663

Constituency Parsing with a Self-Attentive Encoder
Nikita Kitaev and Dan Klein 2675

Pre- and In-Parsing Models for Neural Empty Category Detection
Yufei Chen, Yuanyuan Zhao, Weiwei Sun and Xiaojun Wan 2686

Composing Finite State Transducers on GPUs
Arturo Argueta and David Chiang 2696

Supervised Treebank Conversion: Data and Approaches
Xinzhou Jiang, Zhenghua Li, Bo Zhang, Min Zhang, Sheng Li and Luo Si 2705

Object-oriented Neural Programming (OONP) for Document Understanding
Zhengdong Lu, Xianggen Liu, Haotian Cui, Yukun Yan and Daqi Zheng 2716

Finding syntax in human encephalography with beam search
John Hale, Chris Dyer, Adhiguna Kuncoro and Jonathan Brennan 2726

Learning to Ask Good Questions: Ranking Clarification Questions using Neural Expected Value of Perfect Information
Sudha Rao and Hal Daumé III 2736

Let's do it "again": A First Computational Approach to Detecting Adverbial Presupposition Triggers
Andre Cianflone, Yulan Feng, Jad Kabbara and Jackie Chi Kit Cheung 2746