Platinum Sponsors:

amazon  Tencent  Google  Baidu

Alibaba Group  Facebook  Bloomberg

Samsung

Gold Sponsors:

ebay  Microsoft  IBM Research  SAP

Elsevier  Maluuba  Recruit Institute of Technology  KPMG

Naver  Line  NEC
Preface: General Chair

Welcome to ACL 2017 in Vancouver, Canada! This is the 55th annual meeting of the Association for Computational Linguistics. A tremendous amount of knowledge has been presented at more than half a century’s worth of our conferences. Hopefully, some of it is still relevant now that deep learning has solved language. We are anticipating one of the largest ACL conferences ever. We had a record number of papers submitted to the conference, and a record number of industry partners joining us as sponsors of the conference. We are on track to be one of the best attended ACL conferences to date. I hope that this year’s conference is intellectually stimulating and that you take home many new ideas and techniques that will help extend your own research.

Each year, the ACL conference is organized by a dedicated team of volunteers. Please thank this year’s organizers for their service to the community when you see them at the conference. Without these people, this conference would not happen: Regina Barzilay and Min-Yen Kan (Program Co-Chairs), Priscilla Rasmussen and Anoop Sarkar (Local Organizing Committee), Wei Xu and Jonathan Berant (Workshop Chairs), Maja Popović and Jordan Boyd-Graber (Tutorial Chairs), Wei Lu, Sameer Singh and Margaret Mitchell (Publication Chairs), Heng Ji and Mohit Bansal (Demonstration Chairs), Spandana Gella, Allyson Ettinger, and Matthieu Labeau (Student Research Workshop Organizers), Cecilia Ovedsootter Alm, Mark Dredze, and Marine Carpuat (Faculty Advisors to the Student Research Workshop), Charley Chan (Publicity Chair), Christian Federmann (Conference Handbook Chair), Maryam Siahbani (Student Volunteer Coordinator), and Nitin Madnani (Webmaster and Appmaster).

The organizers have been working for more than a year to put together the conference. Far more than a year in advance, the ACL 2017 Coordinating Committee helped to select the venue and to pick the General Chair and the Program Co-Chairs. This consisted of members from NAACL and ACL executive boards. Representing NAACL we had Hal Daumé III, Michael White, Joel Tetreault, and Emily Bender. Representing ACL we had Pushpak Bhattacharyya, Dragomir Radev, Graeme Hirst, Yejin Choi, and Priscilla Rasmussen. I would like to extend a personal thanks to Graeme and Priscilla who often serve as the ACL’s institutional memory, and who have helped fill in many details along the way.

I would like to extend a special thanks to our Program Co-Chairs, Regina Barzilay and Min-Yen Kan. They documented their work creating the program by running a blog. They used their blog as a platform for engaging the ACL community in many of the decision making processes including soliciting suggestions for the conference’s area chairs and invited speakers. They hosted discussions with Marti Hearst and Joakim Nivre about the value of publishing pre-prints of submitted paper on arXiv and how they relate to double blind reviewing. They even invited several prominent members of our community to provide last-minute writing advice. If you weren’t following the blog in the lead-up to the conference, I highly recommend taking a look through it now. You can find it linked from the ACL 2017 web page.

This year’s program looks like it will be excellent! We owe a huge thank you to Regina Barzilay and Min-Yen Kan. They selected this year’s papers from 1,318 submissions with the help of 44 area chairs and more than 1,200 reviewers. Thanks to Regina, Min, the area chairs, the reviewers and the authors. Beyond the papers, we have talks by luminaries in the field of NLP, including ACL President Joakim Nivre, invited speakers Mirella Lapata and Noah Smith, and the recipient of this year’s Lifetime Achievement Award. We also have an excellent set of workshops and tutorials. On the tutorial day, there will also be a special workshop on Women and Underrepresented Minorities in Natural Language Processing. Thank you to our workshop organizers and tutorial presenters.

This year’s conference features two outreach activities that I would like to highlight. First, on Sunday, July 30, 2017, there will be a workshop on Women and Underrepresented Minorities in Natural Language Processing organized by Libby Barak, Isabelle Augenstein, Chloé Braud, He He, and Margaret Mitchell. The goals of the workshop are to increase awareness of the work women and underrepresented
groups do, support women and underrepresented groups in continuing to pursue their research, and moti-
tivate long-term resources for underrepresented groups within ACL. Second, for the first time ever, ACL
is offering subsidized on-site childcare at the conference hotel. The goal of this is to allow ACL partic-
ipants with children to more readily be able to attend the conference. Since childcare duties often fall
disproportionately on women, our hope is that by having professional childcare on-site that we will allow
more women to participate, and therefore to help promote their careers. My hope is that the childcare
will be continued in future conferences.

I would like to thank our many sponsors for their generous contributions. Our platinum sponsors are Al-
ibaba, Amazon, Apple, Baidu, Bloomberg, Facebook, Google, Samsung and Tencent. Our gold sponsors
are eBay, Elsevier, IBM Research, KPMG, Maluuba, Microsoft, Naver Line, NEC, Recruit Institute of
Technology, and SAP. Our silver sponsors are Adobe, Bosch, CVTE, Duolingo, Huawei, Nuance, Oracle,
and Sogou. Our bronze sponsors are Grammarly, Toutiao, and Yandex. Our supporters include Newsela
and four professional master’s degree programs from Brandeis, Columbia, NYU and the University of
Washington. We would like to acknowledge the generous support of the National Science Foundation
which has awarded a $15,000 grant to the ACL Student Research Workshop. Finally, NVIDIA donated
several Titan X GPU cards for us to raffle off during the conference.

Lastly, I would like to thank everyone else who helped to make this conference a success. Thank you
to our area chairs, our army of reviewers, our workshop organizers, our tutorial presenters, our invited
speakers, and our authors. Best regards to all of you.

Welcome to ACL 2017!

Chris Callison-Burch
General Chair
Welcome to the 55th Annual Meeting of the Association for Computational Linguistics! This year, ACL received 751 long paper submissions and 567 short paper submissions. Of the long papers, 195 were accepted for presentation at ACL — 117 as oral presentations and 78 as poster presentations (25% acceptance rate). 107 short papers were accepted — 34 as oral and 73 as poster presentations (acceptance rate of 18%). In addition, ACL will also feature 21 presentations of papers accepted in the Transactions of the Association for Computational Linguistics (TACL). Including the student research workshop and software demonstrations, the ACL program swells to a massive total of 367 paper presentations on the scientific program, representing the largest ACL program to date.

ACL 2017 will have two distinguished invited speakers: Noah A. Smith (Associate Professor of Computer Science and Engineering at the University of Washington) and Mirella Lapata (Professor in the School of Informatics at the University of Edinburgh). Both are well-renowned for their contributions to the field of computational linguistics and are excellent orators. We are honored that they have accepted our invitation to address the membership at this exciting juncture in our field’s history, addressing key issues in representation learning and multimodal machine translation.

To manage the tremendous growth of our field, we introduced some changes to the conference. With the rotation of the annual meeting to the Americas, we anticipated a heavy load of submissions and early on we decided to have both the long and short paper deadlines merged to reduce reviewing load and to force authors to take a stand on their submissions’ format. The joint deadline allowed us to only load our reviewers once, and also enabled us to have an extended period for more lengthy dialogue among authors, reviewers and area chairs.

In addition, oral presentations were shortened to fourteen (twelve) minutes for long (short) papers, plus time for questions. While this places a greater demand on speakers to be concise, we believe it is worth the effort, allowing far more work to be presented orally. We also took advantage of the many halls available and expanded the number of parallel talks to five during most of the conference sessions.

In keeping with changes introduced in the ACL community from last year, we continued the practice of recognizing outstanding papers at ACL. The 22 outstanding papers (15 long, 7 short, 1.6% of submissions) represent a broad spectrum of exciting contributions and have been specially placed on the final day of the main conference where the program is focused into two parallel sessions of these outstanding contributions. From these, a best paper and a best short paper those will be announced in the awards session on Wednesday afternoon.

Chris has already mentioned our introduction of the chairs’ blog, where we strove to make the selection process of the internal workings of the scientific committee more transparent. We have publicly documented our calls for area chairs, reviewers and accepted papers selection process. Via the blog, we communicated several innovations in the conference organization workflow, of which we would call attention to two key ones here.

In the review process, we pioneered the use of the Toronto Paper Matching System, a topic model based approach to the assignment of reviewers to papers. We hope this decision will spur other program chairs to adopt the system, as increased coverage will better the reviewer/submission matching process, ultimately leading to a higher quality program.

For posterity, we also introduced the usage of hyperlinks in the bibliography reference sections of papers,

1These numbers exclude papers that were not reviewed due to formatting, anonymity, or double submission violations or that were withdrawn prior to review, which was unfortunately a substantial number.

2https://chairs-blog.acl2017.org/
and have worked with the ACL Anthology to ensure that digital object identifiers (DOIs) appear in the footer of each paper. These steps will help broaden the long-term impact of the work that our community has on the scientific world at large.

There are many individuals we wish to thank for their contributions to ACL 2017, some multiple times:

- The 61 area chairs who volunteered for our extra duty. They recruited reviewers, led discussions on each paper, replied to authors’ direct comments to them and carefully assessed each submission. Their input was instrumental in guiding the final decisions on papers and selecting the outstanding papers.

- Our full program committee of BUG hard-working individuals who reviewed the conference’s 1,318 submissions (including secondary reviewers).

- TACL editors-in-chief Mark Johnson, Lillian Lee, and Kristina Toutanova, for coordinating with us on TACL presentations at ACL.

- Noah Smith and Katrin Erk, program co-chairs of ACL 2016 and Ani Nenkova and Owen Rambow, program co-chairs of NAACL 2016, who we consulted several times on short order for help and advice.

- Wei Lu and Sameer Singh, our well-organized publication chairs, with direction and oversight from publication chair mentor Meg Mitchell. Also, Christian Federmann who helped with the local handbook.

- The responsive team at Softconf led by Rich Gerber, who worked quickly to resolve problems and who strove to integrate the use of the Toronto Paper Matching System (TPMS) for our use.

- Priscilla Rasmussen and Anoop Sarkar and the local organization team, especially webmaster Nitin Madnani.

- Christopher Calliston-Burch, our general chair, who kept us coordinated with the rest of the ACL 2017 team and helped us free our time to concentrate on the key duty of organizing the scientific program.

- Key-Sun Choi, Jing Jiang, Graham Neubig, Emily Pitler, and Bonnie Webber who carefully reviewed papers under consideration for best paper recognition.

- Our senior correspondents for the blog, who contributed guest posts and advice for writing and reviewing: Waleed Ammar, Yoav Artzi, Tim Baldwin, Marco Baroni, Claire Cardie, Xavier Carreras, Hal Daumé, Kevin Duh, Chris Dyer, Marti Hearst, Mirella Lapata, Emily M. Bender, Aurélien Max, Kathy McKeown, Ray Mooney, Ani Nenkova, Joakim Nivre, Philip Resnik, and Joel Tetreault. Without them, the participation of the community through the productive comments, and without you the readership, our blog for disseminating information about the decision processes would not have been possible and a success.

We hope that you enjoy ACL 2017 in Vancouver!

ACL 2017 program co-chairs
Regina Barzilay, Massachusetts Institute of Technology
Min-Yen Kan, National University of Singapore
Organizing Committee

General Chair:
Chris Callison-Burch, University of Pennsylvania

Program Co-Chairs:
Regina Barzilay, Massachusetts Institute of Technology
Min-Yen Kan, National University of Singapore

Local Organizing Committee:
Priscilla Rasmussen, ACL
Anoop Sarkar, Simon Fraser University

Workshop Chairs:
Wei Xu, Ohio State University
Jonathan Berant, Tel Aviv University

Tutorial Chairs:
Maja Popović, Humboldt-Universität zu Berlin
Jordan Boyd-Graber, University of Colorado, Boulder

Publication Chairs:
Wei Lu, Singapore University of Technology and Design
Sameer Singh, University of California, Irvine
Margaret Mitchell, Google (Advisory)

Demonstration Chairs:
Heng Ji, Rensselaer Polytechnic Institute
Mohit Bansal, University of North Carolina, Chapel Hill

Student Research Workshop Organizers:
Spandana Gella, University of Edinburgh
Allyson Ettinger, University of Maryland, College Park
Matthieu Labeau, Laboratoire d’Informatique pour la Mécanique et les Sciences de l’Ingénieur (LIMSI)

Faculty Advisors to the Student Research Workshop:
Cecilia Ovesdotter Alm, Rochester Institute of Technology
Mark Dredze, Johns Hopkins University
Marine Carpuat, University of Maryland, College Park
Publicity Chair:
Charley Chan, Bloomberg

Conference Handbook Chair:
Christian Federmann, Microsoft

Student Volunteer Coordinator:
Maryam Siahbani, University of the Fraser Valley

Webmaster and Appmaster:
Nitin Madnani, Educational Testing Service
Program Committee

Program Committee Co-Chairs
Regina Barzilay, Massachusetts Institute of Technology
Min-Yen Kan, National University of Singapore

Area Chairs
Mausam (Information Extraction and NLP Applications Area)
Omri Abend (Multilingual Area)
Eugene Agichtein (Information Extraction and NLP Applications Area)
Ron Artstein (Dialogue and Interactive Systems Area)
Alexandra Balahur (Sentiment Analysis and Opinion Mining Area)
Mohit Bansal (Vision, Robotics and Grounding Area)
Chia-Hui Chang (Information Extraction and NLP Applications Area)
Grzegorz Chrupała (Machine Learning Area)
Mona Diab (Multilingual Area)
Jason Eisner (Phonology, Morphology and Word Segmentation Area)
Manaal Faruqui (Semantics Area)
Raquel Fernandez (Dialogue and Interactive Systems Area)
Karen Fort (Multidisciplinary Area)
Amir Globerson (Machine Learning Area)
Hannaneh Hajishirzi (Semantics Area)
Chiori Hori (Speech Area)
Tommi Jaakkola (Machine Learning Area)
Yangfeng Ji (Discourse and Pragmatics Area)
Jing Jiang (Information Extraction and NLP Applications Area)
Sarvnaz Karimi (Information Extraction and NLP Applications Area)
Anna Korhonen (Semantics Area)
Zornitsa Kozareva (Information Extraction and NLP Applications Area)
Lun-Wei Ku (Sentiment Analysis and Opinion Mining Area)
Nate Kushman (Vision, Robotics and Grounding Area)
Chia-ying Lee (Speech Area)
Oliver Lemon (Dialogue and Interactive Systems Area)
Roger Levy (Cognitive Modeling and Psycholinguistics Area)
Sujian Li (Discourse and Pragmatics Area)
Wenjie Li (Summarization and Generation Area)
Kang Liu (Information Extraction and NLP Applications Area)
Tie-Yan Liu (Information Extraction and NLP Applications Area)
Yang Liu (Machine Translation Area)
Zhiyuan Liu (Social Media Area)
Minh-Thang Luong (Machine Translation Area)
Saif M Mohammad (Sentiment Analysis and Opinion Mining Area)
Alexander M Rush (Summarization and Generation Area)
Haitao Mi (Machine Translation Area)
Alessandro Moschitti (Information Extraction and NLP Applications Area)
Smaranda Muresan (Information Extraction and NLP Applications Area)
Preslav Nakov (Semantics Area)
Graham Neubig (Machine Translation Area)
Primary Reviewers

Reviewers who are acknowledged by the program committee for providing one or more outstanding reviews are marked with “*”.


Juan Antonio Pérez-Ortiz, Ashequl Qadir, Peng Qi, Longhua Qian, Xian Qian, Lu Qin, Long Qiu, Xipeng Qiu, Lizhen Qu, Ariadna Quattoni, Chris Quirk* Alexandre Rademaker, Will Radford, Alessandro Raganato, Afshin Rahimi*, Altaf Rahman, Maya Ramanan, Rohan Ramanathan, A Ramanathan, Arti Ramesh, Gabriela Ramirez-de-la-Rosa, Carlos Ramisch, Anita Ramm, Vivek Kumar Rangarajan Sridhar, Ari Rappoport, Mohammad Sadegh Rasooli, Pushpendre Rastogi, An-
Aris Xanthos, Rui Xia, Yunqing Xia, Bing Xiang, Min Xiao, Tong Xiao, Xinyan Xiao, Boyi Xie, Pengtao Xie, Shasha Xie, Chenyan Xiong, Feiyu Xu, Hua Xu, Ruifeng Xu, Wei Xu, Wenduan Xu, Huichao Xue, Nianwen Xue

Yadollah Yaghoobzadeh, Ichiro Yamada, Bishan Yang, Cheng Yang, Diyi Yang, Grace Hui Yang, Min Yang, Yaqin Yang, Yi Yang, Roman Yangarber, Mark Yatskar, Meliha Yetisgen, Wen-tau Yih, Pengcheng Yin, Wenpeng Yin, Anssi Yli-Jyrä, Dani Yogatama, Naoki Yoshinaga, Bei Yu, Dian Yu, Dianhai Yu, Kai Yu, Liang-Chih Yu, Mo Yu, Zhou Yu, François Yvon


Secondary Reviewers

Naveed Afzal, Yamen Ajjour
Jeremy Barnes, Joost Bastings, Joachim Bingel, Luana Bulat
Iacer Calixto, Lea Canales, Kai Chen, Tongfei Chen, Hao Cheng, Jianpeng Cheng, Yiming Cui
Marco Damonte, Saman Daneshvar, Tobias Domhan, Daxiang Dong, Li Dong
Mohamed Eldesouki
Stefano Faralli, Bin Fu
Srinivasa P. Gadde, Qiaozi Gao, Luca Gilardi, Sujatha Das Gollapalli, J. Manuel Gomez, Stig-Arne Grönroos, Lin Gui
Casper Hansen, Lihong He, Martin Horn
Oana Inel
Gongye Jin
Roman Kern, Vaibhav Kesarwani, Joo-Kyung Kim, Seongchan Kim, Christine Köhn, Santosh Kosgi
Ronja Laarmann-Quante, Egoitz Laparra, Anais Lefeuivre-Halftermeyer, Guanlin Li, Jing Li, Minglei Li, Xiang Li, Xiaolong Li, Chen Liang, Ming Liao, Sijia Liu, Pranay Lohia
Chunpeng Ma, Shuming Ma, Tengfei Ma, Ana Marasovic
Toan Nguyen, Eric Nichols, Sergiu Nisioi
Gözde Özbal
Alexis Palmer, Suraj Pandey, Nikolaos Pappas, José M. Perea-Ortega, Marten Postma
Longhua Qian
Masoud Rouhizadeh Abeer Sarker, Andrew Schneider, Roxane Segers, Pararth Shah, Samiulla Shaikh, Xing Shi, Tomohide Shibata, Samiulla Shiekh, Miikka Silfverberg
Bo Wang*, Boli Wang, Jianxiang Wang, Jingjing Wang, Rui Wang, Shuai Wang, Shuting Wang, Tsung-Hsien Wen, John Wieting
Nan Yang, Yi Yang, Mark Yatskar, Yichun Yin
Sheng Zhang, Kai Zhao, Imed Zitouni
With twin upward trends in the interest in computational linguistics and natural language processing and the size of our annual meeting, ACL has begun the practice of recognizing outstanding papers that represent a select cross-section of the entire field, as nominated by reviewers and vetted by the area chairs and program co-chairs. These papers have been centrally located in the program, on the last day of our meeting, in a more focused two parallel tracks format.

This year, we have nominated 15 long papers and 7 short papers, representing 1.8% of all submissions and approximately 5% of the accepted ACL program. Congratulations, authors!

(in alphabetical order by first author surname)

**Long Papers**

- Jan Buys and Phil Blunsom. *Robust Incremental Neural Semantic Graph Parsing.*
- Daniel Hershcovich, Omri Abend and Ari Rappoport. *A Transition-Based Directed Acyclic Graph Parser for UCCA.*
- Yasuhide Miura, Motoki Taniguchi, Tomoki Taniguchi and Tomoko Ohkuma. *Unifying Text, Metadata, and User Network Representations with a Neural Network for Geolocation Prediction.*
- Ines Rehbein and Josef Ruppenhofer. *Detecting annotation noise in automatically labelled data.*
- Jiwei Tan, Xiaojun Wan and Jianguo Xiao. *Abstractive Document Summarization with a Graph-Based Attentional Neural Model.*
- Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou and Bo Xu. *Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme.*
Short Papers

- Jindřich Libovický and Jindřich Helcl. *Attention Strategies for Multi-Source Sequence-to-Sequence Learning*.
- Keisuke Sakaguchi, Matt Post and Benjamin Van Durme. *Error-repair Dependency Parsing for Ungrammatical Texts*.
Invited Talk: Squashing Computational Linguistics

Noah A. Smith

Paul G. Allen School of Computer Science and Engineering, University of Washington

Abstract

The computational linguistics and natural language processing community is experiencing an episode of deep fascination with representation learning. Like many other presenters at this conference, I will describe new ways to use representation learning in models of natural language. Noting that a data-driven model always assumes a theory (not necessarily a good one), I will argue for the benefits of language-appropriate inductive bias for representation-learning-infused models of language. Such bias often comes in the form of assumptions baked into a model, constraints on an inference algorithm, or linguistic analysis applied to data. Indeed, many decades of research in linguistics (including computational linguistics) put our community in a strong position to identify promising inductive biases. The new models, in turn, may allow us to explore previously unavailable forms of bias, and to produce findings of interest to linguistics. I will focus on new models of documents and of sentential semantic structures, and I will emphasize abstract, reusable components and their assumptions rather than applications.

Biography

Noah Smith is an Associate Professor in the Paul G. Allen School of Computer Science and Engineering at the University of Washington. Previously, he was an Associate Professor in the School of Computer Science at Carnegie Mellon University. He received his Ph.D. in Computer Science from Johns Hopkins University and his B.S. in Computer Science and B.A. in Linguistics from the University of Maryland. His research spans many topics in natural language processing, machine learning, and computational social science. He has served on the editorial boards of CL, JAIR, and TACL, as the secretary-treasurer of SIGDAT (2012–2015), and as program co-chair of ACL 2016. Alumni of his research group, Noah’s ARK, are international leaders in NLP in academia and industry. Smith’s work has been recognized with a UW Innovation award, a Finmeccanica career development chair at CMU, an NSF CAREER award, a Hertz Foundation graduate fellowship, numerous best paper nominations and awards, and coverage by NPR, BBC, CBC, the New York Times, the Washington Post, and Time.
Invited Talk: Translating from Multiple Modalities to Text and Back

Mirella Lapata

Professor, School of Informatics, University of Edinburgh

Abstract

Recent years have witnessed the development of a wide range of computational tools that process and generate natural language text. Many of these have become familiar to mainstream computer users in the form of web search, question answering, sentiment analysis, and notably machine translation. The accessibility of the web could be further enhanced with applications that not only translate between different languages (e.g., from English to French) but also within the same language, between different modalities, or different data formats. The web is rife with non-linguistic data (e.g., video, images, source code) that cannot be indexed or searched since most retrieval tools operate over textual data.

In this talk I will argue that in order to render electronic data more accessible to individuals and computers alike, new types of translation models need to be developed. I will focus on three examples, text simplification, source code generation, and movie summarization. I will illustrate how recent advances in deep learning can be extended in order to induce general representations for different modalities and learn how to translate between these and natural language.

Biography

Mirella Lapata is professor of natural language processing in the School of Informatics at the University of Edinburgh. Her research focuses on getting computers to understand, reason with, and generate. She is as an associate editor of the Journal of Artificial Intelligence Research and has served on the editorial boards of Transactions of the ACL and Computational Linguistics. She was the first recipient of the Karen Sparck Jones award of the British Computer Society, recognizing key contributions to NLP and information retrieval. She received two EMNLP best paper awards and currently holds a prestigious Consolidator Grant from the European Research Council.
# Table of Contents

**Adversarial Multi-task Learning for Text Classification**  
Pengfei Liu, Xipeng Qiu and Xuanjing Huang  
1

**Neural End-to-End Learning for Computational Argumentation Mining**  
Steffen Eger, Johannes Daxenberger and Iryna Gurevych  
11

**Neural Symbolic Machines: Learning Semantic Parsers on Freebase with Weak Supervision**  
Chen Liang, Jonathan Berant, Quoc Le, Kenneth D. Forbus and Ni Lao  
23

**Neural Relation Extraction with Multi-lingual Attention**  
Yankai Lin, Zhiyuan Liu and Maosong Sun  
34

**Learning Structured Natural Language Representations for Semantic Parsing**  
Jianpeng Cheng, Siva Reddy, Vijay Saraswat and Mirella Lapata  
44

**Morph-fitting: Fine-Tuning Word Vector Spaces with Simple Language-Specific Rules**  
Ivan Vulić, Nikola Mrkšić, Roi Reichart, Diarmuid Ó Séaghdha, Steve Young and Anna Korhonen  
56

**Skip-Gram - Zipf + Uniform = Vector Additivity**  
Alex Gittens, Dimitris Achlioptas and Michael W. Mahoney  
69

**The State of the Art in Semantic Representation**  
Omri Abend and Ari Rappoport  
77

**Joint Learning for Event Coreference Resolution**  
Jing Lu and Vincent Ng  
90

**Generating and Exploiting Large-scale Pseudo Training Data for Zero Pronoun Resolution**  
Ting Liu, Yiming Cui, Qingyu Yin, Wei-Nan Zhang, Shijin Wang and Guoping Hu  
102

**Discourse Mode Identification in Essays**  
Wei Song, Dong Wang, Ruiji Fu, Lizhen Liu, Ting Liu and Guoping Hu  
112

**A Convolutional Encoder Model for Neural Machine Translation**  
Jonas Gehring, Michael Auli, David Grangier and Yann Dauphin  
123

**Deep Neural Machine Translation with Linear Associative Unit**  
Mingxuan Wang, Zhengdong Lu, Jie Zhou and Qun Liu  
136

**Neural AMR: Sequence-to-Sequence Models for Parsing and Generation**  
Ioannis Konstas, Srinivasan Iyer, Mark Yatskar, Yejin Choi and Luke Zettlemoyer  
146

**Program Induction by Rationale Generation: Learning to Solve and Explain Algebraic Word Problems**  
Wang Ling, Dani Yogatama, Chris Dyer and Phil Blunsom  
158

**Automatically Generating Rhythmic Verse with Neural Networks**  
Jack Hopkins and Douwe Kiela  
168

**Creating Training Corpora for NLG Micro-Planners**  
Claire Gardent, Anastasia Shimorina, Shashi Narayan and Laura Perez-Beltrachini  
179
Gated Self-Matching Networks for Reading Comprehension and Question Answering
Wenhui Wang, Nan Yang, Furu Wei, Baobao Chang and Ming Zhou ................................. 189

Generating Natural Answers by Incorporating Copying and Retrieving Mechanisms in Sequence-to-Sequence Learning
Shizhu He, Cao Liu, Kang Liu and Jun Zhao ................................................................. 199

Coarse-to-Fine Question Answering for Long Documents
Eunsol Choi, Daniel Hewlett, Jakob Uszkoreit, Illia Polosukhin, Alexandre Lacoste and Jonathan Berant ................................................................. 209

An End-to-End Model for Question Answering over Knowledge Base with Cross-Attention Combining Global Knowledge
Yanchao Hao, Yuanzhe Zhang, Kang Liu, Shizhu He, Zhanyi Liu, Hua Wu and Jun Zhao ...... 221

Translating Neuraless
Jacob Andreas, Anca Dragan and Dan Klein ................................................................. 232

Obtaining referential word meanings from visual and distributional information: Experiments on object naming
Sina Zarrieß and David Schlangen ................................................................. 243

FOIL it! Find One mismatch between Image and Language caption
Ravi Shekhar, Sandro Pezzelle, Yauhen Klimovich, Aurélie Herbelot, Moin Nabi, Enver Sangineto and Raffaella Bernardi ................................................................. 255

Verb Physics: Relative Physical Knowledge of Actions and Objects
Maxwell Forbes and Yejin Choi ................................................................. 266

A* CCG Parsing with a Supertag and Dependency Factored Model
Masashi Yoshikawa, Hiroshi Noji and Yuji Matsumoto ................................................................. 277

A Full Non-Monotonic Transition System for Unrestricted Non-Projective Parsing
Daniel Fernández-González and Carlos Gómez-Rodríguez ................................................................. 288

Aggregating and Predicting Sequence Labels from Crowd Annotations
An Thanh Nguyen, Byron Wallace, Junyi Jessy Li, Ani Nenkova and Matthew Lease ...... 299

Multi-space Variational Encoder-Decoders for Semi-supervised Labeled Sequence Transduction
Chunting Zhou and Graham Neubig ................................................................. 310

Scalable Bayesian Learning of Recurrent Neural Networks for Language Modeling
Zhe Gan, Chunyuan Li, Changyou Chen, Yunchen Pu, Qinliang Su and Lawrence Carin ...... 321

Learning attention for historical text normalization by learning to pronounce
Marcel Bollmann, Joachim Bingel and Anders Søgaard ................................................................. 332

Deep Learning in Semantic Kernel Spaces
Danilo Croce, Simone Filice, Giuseppe Castellucci and Roberto Basili ................................................................. 345

Topically Driven Neural Language Model
Jey Han Lau, Timothy Baldwin and Trevor Cohn ................................................................. 355

Handling Cold-Start Problem in Review Spam Detection by Jointly Embedding Texts and Behaviors
Xuepeng Wang, Kang Liu and Jun Zhao ................................................................. 366
Learning Cognitive Features from Gaze Data for Sentiment and Sarcasm Classification using Convolutional Neural Network
Abhijit Mishra, Kuntal Dey and Pushpak Bhattacharyya ............................... 377

An Unsupervised Neural Attention Model for Aspect Extraction
Ruidan He, Wee Sun Lee, Hwee Tou Ng and Daniel Dahlmeier .......................... 388

Other Topics You May Also Agree or Disagree: Modeling Inter-Topic Preferences using Tweets and Matrix Factorization
Akira Sasaki, Kazuaki Hanawa, Naoaki Okazaki and Kentaro Inui ....................... 398

Automatically Labeled Data Generation for Large Scale Event Extraction
Yubo Chen, Shulin Liu, Xiang Zhang, Kang Liu and Jun Zhao .............................. 409

Time Expression Analysis and Recognition Using Syntactic Token Types and General Heuristic Rules
Xiaoshi Zhong, Aixin Sun and Erik Cambria ..................................................... 420

Learning with Noise: Enhance Distantly Supervised Relation Extraction with Dynamic Transition Matrix
Bingfeng Luo, Yansong Feng, Zheng Wang, Zhanxing Zhu, Songfang Huang, Rui Yan and Dongyan Zhao ................................................................. 430

A Syntactic Neural Model for General-Purpose Code Generation
Pengcheng Yin and Graham Neubig ................................................................. 440

Learning bilingual word embeddings with (almost) no bilingual data
Mikel Artetxe, Gorka Labaka and Eneko Agirre ................................................. 451

Abstract Meaning Representation Parsing using LSTM Recurrent Neural Networks
William Foland and James H. Martin ............................................................... 463

Deep Semantic Role Labeling: What Works and What’s Next
Luheng He, Kenton Lee, Mike Lewis and Luke Zettlemoyer ............................... 473

Towards End-to-End Reinforcement Learning of Dialogue Agents for Information Access
Bhuwan Dhingra, Lihong Li, Xiujuin Li, Jianfeng Gao, Yun-Nung Chen, Faisal Ahmed and Li Deng 484

Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-Based Chatbots
Yu Wu, Wei Wu, Chen Xing, Ming Zhou and Zhoujun Li ................................. 496

Learning Word-Like Units from Joint Audio-Visual Analysis
David Harwath and James Glass ................................................................. 506

Joint CTC/attention decoding for end-to-end speech recognition
Takaaki Hori, Shinji Watanabe and John Hershey ........................................ 518

Found in Translation: Reconstructing Phylogenetic Language Trees from Translations
Ella Rabinovich, Noam Ordan and Shuly Wintner ......................................... 530

Predicting Native Language from Gaze
Yevgeni Berzak, Chie Nakamura, Suzanne Flynn and Boris Katz ....................... 541

MORSE: Semantic-ally Drive-n MORpheme SEgment-er
Tarek Sakakini, Suma Bhat and Pramod Viswanath ...................................... 552
Deep Pyramid Convolutional Neural Networks for Text Categorization
Rie Johnson and Tong Zhang ................................................................. 562

Improved Neural Relation Detection for Knowledge Base Question Answering
Mo Yu, Wenpeng Yin, Kazi Saidul Hasan, Cicero dos Santos, Bing Xiang and Bowen Zhou ..... 571

Deep Keyphrase Generation
Rui Meng, Sanqiang Zhao, Shuguang Han, Daqing He, Peter Brusilovsky and Yu Chi ............. 582

Attention-over-Attention Neural Networks for Reading Comprehension
Yiming Cui, Zhipeng Chen, Si Wei, Shijin Wang, Ting Liu and Guoping Hu .......................... 593

Alignment at Work: Using Language to Distinguish the Internalization and Self-Regulation Components of Cultural Fit in Organizations
Gabriel Doyle, Amir Goldberg, Sameer Srivastava and Michael Frank .............................. 603

Representations of language in a model of visually grounded speech signal
Grzegorz Chrupała, Lieke Gelderloos and Afra Alishahi .................................................. 613

Spectral Analysis of Information Density in Dialogue Predicts Collaborative Task Performance
Yang Xu and David Reitter ................................................................. 623

Affect-LM: A Neural Language Model for Customizable Affective Text Generation
Sayan Ghosh, Mathieu Chollet, Eugene Laksana, Louis-Philippe Morency and Stefan Scherer 634

Domain Attention with an Ensemble of Experts
Young-Bum Kim, Karl Stratos and Dongchan Kim .......................................................... 643

Learning Discourse-level Diversity for Neural Dialog Models using Conditional Variational Autoencoders
Tiancheng Zhao, Ran Zhao and Maxine Eskenazi ............................................................. 654

Hybrid Code Networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning
Jason D Williams, Kavosh Asadi and Geoffrey Zweig ....................................................... 665

Generating Contrastive Referring Expressions
Martin Villalba, Christoph Teichmann and Alexander Koller ........................................... 678

Modeling Source Syntax for Neural Machine Translation
Junhui Li, Deyi Xiong, Zhaopeng Tu, Muhua Zhu, Min Zhang and Guodong Zhou ............... 688

Sequence-to-Dependency Neural Machine Translation
Shuangzi Wu, Dongdong Zhang, Nan Yang, Mu Li and Ming Zhou .................................... 698

Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning
Jing Ma, Wei Gao and Kam-Fai Wong ............................................................................. 708

EmoNet: Fine-Grained Emotion Detection with Gated Recurrent Neural Networks
Muhammad Abdul-Mageed and Lyle Ungar ..................................................................... 718

Beyond Binary Labels: Political Ideology Prediction of Twitter Users
Daniel Preoţiuc-Pietro, Ye Liu, Daniel Hopkins and Lyle Ungar ....................................... 729

xxviii
Leveraging Behavioral and Social Information for Weakly Supervised Collective Classification of Political Discourse on Twitter
Kristen Johnson, Di Jin and Dan Goldwasser ................................................................. 741

A Nested Attention Neural Hybrid Model for Grammatical Error Correction
Jianshu Ji, Qinlong Wang, Kristina Toutanova, Yongen Gong, Steven Truong and Jianfeng Gao 753

TextFlow: A Text Similarity Measure based on Continuous Sequences
Yassine Mrabet, Halil Kilicoglu and Dina Demner-Fushman ........................................ 763

Friendships, Rivalries, and Trysts: Characterizing Relations between Ideas in Texts
Chenhao Tan, Dallas Card and Noah A. Smith ................................................................. 773

Polish evaluation dataset for compositional distributional semantics models
Alina Wróblewska and Katarzyna Krasnowska-Kieraś .................................................... 784

Automatic Annotation and Evaluation of Error Types for Grammatical Error Correction
Christopher Bryant, Mariano Felice and Ted Briscoe ..................................................... 793

Evaluation Metrics for Machine Reading Comprehension: Prerequisite Skills and Readability
Saku Sugawara, Yusuke Kido, Hikaru Yokono and Akiko Aizawa .................................. 806

A Minimal Span-Based Neural Constituency Parser
Mitchell Stern, Jacob Andreas and Dan Klein ............................................................... 818

Semantic Dependency Parsing via Book Embedding
Weiwei Sun, Junjie Cao and Xiaojun Wan ................................................................. 828

Neural Word Segmentation with Rich Pretraining
Jie Yang, Yue Zhang and Fei Dong ........................................................................ 839

Neural Machine Translation via Binary Code Prediction
Yusuke Oda, Philip Arthur, Graham Neubig, Koichiro Yoshino and Satoshi Nakamura .... 850

What do Neural Machine Translation Models Learn about Morphology?
Yonatan Belinkov, Nadir Durran, Fahim Dalvi, Hassan Sajjad and James Glass ............. 861

Context-Dependent Sentiment Analysis in User-Generated Videos
Soujanya Poria, Erik Cambria, Devamanyu Hazarika, Navonil Majumder, Amir Zadeh and Louis-Philippe Morency .................................................. 873

A Multidimensional Lexicon for Interpersonal Stancetaking
Umashanthi Pavalanathan, Jim Fitzpatrick, Scott Kiesling and Jacob Eisenstein ........... 884

Tandem Anchoring: a Multiword Anchor Approach for Interactive Topic Modeling
Jeffrey Lund, Connor Cook, Kevin Seppi and Jordan Boyd-Graber ............................... 896

Apples to Apples: Learning Semantics of Common Entities Through a Novel Comprehension Task
Omid Bakhshandeh and James Allen ........................................................................ 906

Going out on a limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees
Arzoo Katiyar and Claire Cardie ........................................................................ 917

Naturalizing a Programming Language via Interactive Learning
Sida I. Wang, Samuel Ginn, Percy Liang and Christopher D. Manning ...................... 929

xxix
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic Word Clusters Using Signed Spectral Clustering</td>
<td>Joao Sedoc, Jean Gallier, Dean Foster and Lyle Ungar</td>
<td>939</td>
</tr>
<tr>
<td>An Interpretable Knowledge Transfer Model for Knowledge Base Completion</td>
<td>Qizhe Xie, Xuezhe Ma, Zihang Dai and Eduard Hovy</td>
<td>950</td>
</tr>
<tr>
<td>Joint Modeling of Content and Discourse Relations in Dialogues</td>
<td>Kechen Qin, Lu Wang and Joseph Kim</td>
<td>974</td>
</tr>
<tr>
<td>Argument Mining with Structured SVMs and RNNs</td>
<td>Vlad Niculae, Joonsuk Park and Claire Cardie</td>
<td>985</td>
</tr>
<tr>
<td>Neural Discourse Structure for Text Categorization</td>
<td>Yangfeng Ji and Noah A. Smith</td>
<td>996</td>
</tr>
<tr>
<td>Adversarial Connective-exploiting Networks for Implicit Discourse Relation Classification</td>
<td>Lianhui Qin, Zhisong Zhang, Hai Zhao, Zhiting Hu and Eric Xing</td>
<td>1006</td>
</tr>
<tr>
<td>Don’t understand a measure? Learn it: Structured Prediction for Coreference Resolution optimizing its measures</td>
<td>Iryna Haponchyk and Alessandro Moschitti</td>
<td>1018</td>
</tr>
<tr>
<td>Bayesian Modeling of Lexical Resources for Low-Resource Settings</td>
<td>Nicholas Andrews, Mark Dredze, Benjamin Van Durme and Jason Eisner</td>
<td>1029</td>
</tr>
<tr>
<td>Semi-Supervised QA with Generative Domain-Adaptive Nets</td>
<td>Zhilin Yang, Junjie Hu, Ruslan Salakhutdinov and William Cohen</td>
<td>1040</td>
</tr>
<tr>
<td>From Language to Programs: Bridging Reinforcement Learning and Maximum Marginal Likelihood</td>
<td>Kelvin Guu, Panupong Pasupat, Evan Liu and Percy Liang</td>
<td>1051</td>
</tr>
<tr>
<td>Diversity driven attention model for query-based abstractive summarization</td>
<td>Preksha Nema, Mitesh M. Khapra, Anirban Laha and Balaraman Ravindran</td>
<td>1063</td>
</tr>
<tr>
<td>Get To The Point: Summarization with Pointer-Generator Networks</td>
<td>Abigail See, Peter J. Liu and Christopher D. Manning</td>
<td>1073</td>
</tr>
<tr>
<td>Supervised Learning of Automatic Pyramid for Optimization-Based Multi-Document Summarization</td>
<td>Maxime Peyrard and Judith Eckle-Kohler</td>
<td>1084</td>
</tr>
<tr>
<td>Selective Encoding for Abstractive Sentence Summarization</td>
<td>Qingyu Zhou, Nan Yang, Furu Wei and Ming Zhou</td>
<td>1095</td>
</tr>
<tr>
<td>PositionRank: An Unsupervised Approach to Keyphrase Extraction from Scholarly Documents</td>
<td>Corina Florescu and Cornelia Caragea</td>
<td>1105</td>
</tr>
<tr>
<td>Towards an Automatic Turing Test: Learning to Evaluate Dialogue Responses</td>
<td>Ryan Lowe, Michael Noseworthy, Iulian Vlad Serban, Nicolas Angelard-Gontier, Yoshua Bengio and Joelle Pineau</td>
<td>1116</td>
</tr>
<tr>
<td>A Transition-Based Directed Acyclic Graph Parser for UCCA</td>
<td>Daniel Hershcovich, Omri Abend and Ari Rappoport</td>
<td>1127</td>
</tr>
</tbody>
</table>

xxx
<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract Syntax Networks for Code Generation and Semantic Parsing</td>
<td>1139</td>
</tr>
<tr>
<td>Maxim Rabinovich, Mitchell Stern and Dan Klein</td>
<td></td>
</tr>
<tr>
<td>Visualizing and Understanding Neural Machine Translation</td>
<td>1150</td>
</tr>
<tr>
<td>Yanzhuo Ding, Yang Liu, Huanbo Luan and Maosong Sun</td>
<td></td>
</tr>
<tr>
<td>Detecting annotation noise in automatically labelled data</td>
<td>1160</td>
</tr>
<tr>
<td>Ines Rehbein and Josef Ruppenhofer</td>
<td></td>
</tr>
<tr>
<td>Abstractive Document Summarization with a Graph-Based Attentional Neural Model</td>
<td>1171</td>
</tr>
<tr>
<td>Jiwei Tan, Xiaojun Wan and Jianguo Xiao</td>
<td></td>
</tr>
<tr>
<td>Probabilistic Typology: Deep Generative Models of Vowel Inventories</td>
<td>1182</td>
</tr>
<tr>
<td>Ryan Cotterell and Jason Eisner</td>
<td></td>
</tr>
<tr>
<td>Adversarial Multi-Criteria Learning for Chinese Word Segmentation</td>
<td>1193</td>
</tr>
<tr>
<td>Xinchi Chen, Zhan Shi, Xipeng Qiu and Xuanjing Huang</td>
<td></td>
</tr>
<tr>
<td>Neural Joint Model for Transition-based Chinese Syntactic Analysis</td>
<td>1204</td>
</tr>
<tr>
<td>Shuhei Kurita, Daisuke Kawahara and Sadao Kurohashi</td>
<td></td>
</tr>
<tr>
<td>Robust Incremental Neural Semantic Graph Parsing</td>
<td>1215</td>
</tr>
<tr>
<td>Jan Buyss and Phil Blunsom</td>
<td></td>
</tr>
<tr>
<td>Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme</td>
<td>1227</td>
</tr>
<tr>
<td>Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou and Bo Xu.</td>
<td></td>
</tr>
<tr>
<td>A Local Detection Approach for Named Entity Recognition and Mention Detection</td>
<td>1237</td>
</tr>
<tr>
<td>Mingbin Xu, Hui Jiang and Sedtawut Watcharawittayakul</td>
<td></td>
</tr>
<tr>
<td>Vancouver Welcomes You! Minimalist Location Metonymy Resolution</td>
<td>1248</td>
</tr>
<tr>
<td>Milan Gritta, Mohammad Taher Pilehvar, Nut Limsoopatham and Nigel Collier</td>
<td></td>
</tr>
<tr>
<td>Unifying Text, Metadata, and User Network Representations with a Neural Network for Geolocation Prediction</td>
<td>1260</td>
</tr>
<tr>
<td>Yasuhide Miura, Motokii Taniguchi, Tomoki Taniguchi and Tomoko Okhuma</td>
<td></td>
</tr>
<tr>
<td>Multi-Task Video Captioning with Video and Entailment Generation</td>
<td>1273</td>
</tr>
<tr>
<td>Ramakanth Pasunuru and Mohit Bansal</td>
<td></td>
</tr>
<tr>
<td>Enriching Complex Networks with Word Embeddings for Detecting Mild Cognitive Impairment from Speech Transcripts</td>
<td>1284</td>
</tr>
<tr>
<td>Leandro Santos, Edilson Anselmo Corrêa Júnior, Osvaldo Oliveira Jr, Diego Amancio, Letícia Mansur and Sandra Aluísio</td>
<td></td>
</tr>
<tr>
<td>Adversarial Adaptation of Synthetic or Stale Data</td>
<td>1297</td>
</tr>
<tr>
<td>Young-Bum Kim, Karl Stratos and Dongchan Kim</td>
<td></td>
</tr>
<tr>
<td>Chat Detection in an Intelligent Assistant: Combining Task-oriented and Non-task-oriented Spoken Dialogue Systems</td>
<td>1308</td>
</tr>
<tr>
<td>Satoshi Akasaki and Nobuhiro Kaji</td>
<td></td>
</tr>
<tr>
<td>A Neural Local Coherence Model</td>
<td>1320</td>
</tr>
<tr>
<td>Dat Tien Nguyen and Shafiq Joty</td>
<td></td>
</tr>
</tbody>
</table>

xxxi
Data-Driven Broad-Coverage Grammars for Opinionated Natural Language Generation (ONLG)
     Tomer Cagan, Stefan L. Frank and Reut Tsarfaty .................................................. 1331

Learning to Ask: Neural Question Generation for Reading Comprehension
     Xinya Du, Junru Shao and Claire Cardie .............................................................. 1342

Joint Optimization of User-desired Content in Multi-document Summaries by Learning from User Feedback
     Avinesh PVS and Christian M. Meyer ................................................................. 1353

Flexible and Creative Chinese Poetry Generation Using Neural Memory
     Jiyuan Zhang, Yang Feng, Dong Wang, Yang Wang, Andrew Abel, Shiyoue Zhang and Andi Zhang 1364

Learning to Generate Market Comments from Stock Prices
     Soichiro Murakami, Akihiko Watanabe, Akira Miyazawa, Keiichi Goshima, Toshihiko Yanase, Hiroya Takamura and Yusuke Miyao ................................................................. 1374

Can Syntax Help? Improving an LSTM-based Sentence Compression Model for New Domains
     Liangguo Wang, Jing Jiang, Hai Leong Chieu, Chen Hui Ong, Dandan Song and Lejian Liao 1385

Transductive Non-linear Learning for Chinese Hypernym Prediction
     Chengyu Wang, Junichi Yan, Aoying Zhou and Xiaofeng He ................................. 1394

A Constituent-Centric Neural Architecture for Reading Comprehension
     Pengtao Xie and Eric Xing ..................................................................................... 1405

Cross-lingual Distillation for Text Classification
     Ruochen Xu and Yiming Yang .............................................................................. 1415

Understanding and Predicting Empathic Behavior in Counseling Therapy
     Verónica Pérez-Rosas, Rada Mihalcea, Kenneth Resnicow, Satinder Singh and Lawrence An. 1426

Leveraging Knowledge Bases in LSTMs for Improving Machine Reading
     Bishan Yang and Tom Mitchell .............................................................................. 1436

Prerequisite Relation Learning for Concepts in MOOCs
     Liangming Pan, Chengjiang Li, Juanzi Li and Jie Tang .............................................. 1447

Unsupervised Text Segmentation Based on Native Language Characteristics
     Shervin Malmasi, Mark Dras, Mark Johnson, Lan Du and Magdalena Wolska .......... 1457

Weakly Supervised Cross-Lingual Named Entity Recognition via Effective Annotation and Representation Projection
     Jian Ni, Georgiana Dinu and Radu Florian .............................................................. 1470

Context Sensitive Lemmatization Using Two Successive Bidirectional Gated Recurrent Networks
     Abhisek Chakrabarty, Onkar Arun Pandit and Utpal Garain .................................. 1481

Learning to Create and Reuse Words in Open-Vocabulary Neural Language Modeling
     Kazuya Kawakami, Chris Dyer and Phil Blunsom ................................................... 1492

Bandit Structured Prediction for Neural Sequence-to-Sequence Learning
     Julia Kreutzer, Artem Sokolov and Stefan Riezler ............................................... 1503
Volatility Prediction using Financial Disclosures Sentiments with Word Embedding-based IR Models
Navid Rekabsaz, Mihai Lupu, Artem Baklanov, Alexander Dür, Linda Andersson and Allan Hanbury

CANE: Context-Aware Network Embedding for Relation Modeling
Cunchao Tu, Han Liu, Zhiyuan Liu and Maosong Sun

Universal Dependencies Parsing for Colloquial Singaporean English
Hongmin Wang, Yue Zhang, Guang Yong Leonard Chan, Jie Yang and Hai Leong Chieu

Generic Axiomatization of Families of Noncrossing Graphs in Dependency Parsing
Anssi Yli-Jyrä and Carlos Gómez-Rodríguez

Semi-supervised sequence tagging with bidirectional language models
Matthew Peters, Waleed Ammar, Chandra Bhagavatula and Russell Power

Learning Symmetric Collaborative Dialogue Agents with Dynamic Knowledge Graph Embeddings
He He, Anusha Balakrishnan, Mihail Eric and Percy Liang

Neural Belief Tracker: Data-Driven Dialogue State Tracking
Nikola Mrkšić, Diarmuid O Séaghdha, Tsung-Hsien Wen, Blaise Thomson and Steve Young

Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms
Shulin Liu, Yubo Chen, Kang Liu and Jun Zhao

Topical Coherence in LDA-based Models through Induced Segmentation
Hesam Amoualian, Wei Lu, Eric Gaussier, Georgios Balikas, Massih R Amini and Marianne Clausel

Jointly Extracting Relations with Class Ties via Effective Deep Ranking
Hai Ye, Wenhan Chao, Zhunchen Luo and Zhoujun Li

Search-based Neural Structured Learning for Sequential Question Answering
Mohit Iyyer, Wen-tau Yih and Ming-Wei Chang

Gated-Attention Readers for Text Comprehension
Bhuwan Dhingra, Hanxiao Liu, Zhilin Yang, William Cohen and Ruslan Salakhutdinov

Determining Gains Acquired from Word Embedding Quantitatively Using Discrete Distribution Clustering
Jianbo Ye, Yanran Li, Zhaohui Wu, James Z. Wang, Wenjie Li and Jia Li

Towards a Seamless Integration of Word Senses into Downstream NLP Applications
Mohammad Taher Pilehvar, Jose Camacho-Collados, Roberto Navigli and Nigel Collier

Reading Wikipedia to Answer Open-Domain Questions
Danqi Chen, Adam Fisch, Jason Weston and Antoine Bordes

Learning to Skim Text
Adams Wei Yu, Hongrae Lee and Quoc Le

An Algebra for Feature Extraction
Vivek Srikumar

xxxiv
Chunk-based Decoder for Neural Machine Translation
Shonosuke Ishiwatari, Jingtao Yao, Shujie Liu, Mu Li, Ming Zhou, Naoki Yoshinaga, Masaru Kitsuregawa and Weijia Jia ................................................................. 1901

Doubly-Attentive Decoder for Multi-modal Neural Machine Translation
Iacer Calixto, Qun Liu and Nick Campbell ...................................................... 1913

A Teacher-Student Framework for Zero-Resource Neural Machine Translation
Yun Chen, Yang Liu, Yong Cheng and Victor O.K. Li ....................................... 1925

Improved Neural Machine Translation with a Syntax-Aware Encoder and Decoder
Huadong Chen, Shujian Huang, David Chiang and Jiajun Chen .............................. 1936

Cross-lingual Name Tagging and Linking for 282 Languages
Xiaoman Pan, Boliang Zhang, Jonathan May, Joel Nothman, Kevin Knight and Heng Ji 1946

Adversarial Training for Unsupervised Bilingual Lexicon Induction
Meng Zhang, Yang Liu, Huanbo Luan and Maosong Sun ..................................... 1959

Estimating Code-Switching on Twitter with a Novel Generalized Word-Level Language Detection Technique
Shruti Rijhwani, Royal Sequiera, Monojit Choudhury, Kalika Bali and Chandra Shekhar Maddila 1971

Using Global Constraints and Reranking to Improve Cognates Detection
Michael Bloodgood and Benjamin Strauss ...................................................... 1983

One-Shot Neural Cross-Lingual Transfer for Paradigm Completion
Katharina Kann, Ryan Cotterell and Hinrich Schütze ....................................... 1993

Morphological Inflection Generation with Hard Monotonic Attention
Roee Aharoni and Yoav Goldberg ................................................................. 2004

From Characters to Words to in Between: Do We Capture Morphology?
Clara Vania and Adam Lopez ............................................................................ 2016

Riemannian Optimization for Skip-Gram Negative Sampling
Alexander Fonarev, Oleksii Grinchuk, Gleb Gusev, Pavel Serdyukov and Ivan Oseledets 2028

Deep Multitask Learning for Semantic Dependency Parsing
Hao Peng, Sam Thomson and Noah A. Smith .................................................. 2037

Improved Word Representation Learning with Sememes
Yilin Niu, Ruobing Xie, Zhiyuan Liu and Maosong Sun ..................................... 2049

Learning Character-level Compositionality with Visual Features
Frederick Liu, Han Lu, Chieh Lo and Graham Neubig ....................................... 2059

A Progressive Learning Approach to Chinese SRL Using Heterogeneous Data
Qiaolin Xia, Lei Sha, Baobao Chang and Zhifang Sui ....................................... 2069

Revisiting Recurrent Networks for Paraphrastic Sentence Embeddings
John Wieting and Kevin Gimpel ................................................................. 2078

Ontology-Aware Token Embeddings for Prepositional Phrase Attachment
Pradeep Dasigi, Waleed Ammar, Chris Dyer and Eduard Hovy ............................. 2089

xxxv
Identifying 1950s American Jazz Musicians: Fine-Grained IsA Extraction via Modifier Composition
Ellie Pavlick and Marius Pasca ................................................................. 2099

Parsing to 1-Endpoint-Crossing, Pagenumber-2 Graphs
Junjie Cao, Sheng Huang, Weiwei Sun and Xiaojun Wan ......................... 2110

Semi-supervised Multitask Learning for Sequence Labeling
Marek Rei .................................................................................................... 2121

Semantic Parsing of Pre-university Math Problems
Takuya Matsuzaki, Takumi Ito, Hidenao Iwane, Hirokazu Anai and Noriko H. Arai .... 2131
Conference Program

Monday, July 31st

10:30–11:45  Session 1A: Information Extraction 1 (NN)

10:30–10:48  Adversarial Multi-task Learning for Text Classification
Pengfei Liu, Xipeng Qiu and Xuanjing Huang

10:49–11:07  Neural End-to-End Learning for Computational Argumentation Mining
Steffen Eger, Johannes Daxenberger and Iryna Gurevych

11:08–11:26  Neural Symbolic Machines: Learning Semantic Parsers on Freebase with Weak Supervision
Chen Liang, Jonathan Berant, Quoc Le, Kenneth D. Forbus and Ni Lao

11:27–11:45  Neural Relation Extraction with Multi-lingual Attention
Yankai Lin, Zhiyuan Liu and Maosong Sun

Monday, July 31st

10:30–11:45  Session 1B: Semantics 1

10:30–10:48  Learning Structured Natural Language Representations for Semantic Parsing
Jianpeng Cheng, Siva Reddy, Vijay Saraswat and Mirella Lapata

10:49–11:07  Morph-fitting: Fine-Tuning Word Vector Spaces with Simple Language-Specific Rules
Ivan Vulić, Nikola Mrkšić, Roi Reichart, Diarmuid Ó Séaghdha, Steve Young and Anna Korhonen

11:08–11:26  Skip-Gram - Zipf + Uniform = Vector Additivity
Alex Gittens, Dimitris Achlioptas and Michael W. Mahoney

11:27–11:45  The State of the Art in Semantic Representation
Omri Abend and Ari Rappoport
Monday, July 31st

10:30–11:26  Session 1C: Discourse 1

10:30–10:48  Joint Learning for Event Coreference Resolution
Jing Lu and Vincent Ng

10:49–11:07  Generating and Exploiting Large-scale Pseudo Training Data for Zero Pronoun Resolution
Ting Liu, Yiming Cui, Qingyu Yin, Wei-Nan Zhang, Shijin Wang and Guoping Hu

11:08–11:26  Discourse Mode Identification in Essays
Wei Song, Dong Wang, Ruiji Fu, Lizhen Liu, Ting Liu and Guoping Hu

Monday, July 31st

10:30–11:07  Session 1D: Machine Translation 1

10:30–10:48  A Convolutional Encoder Model for Neural Machine Translation
Jonas Gehring, Michael Auli, David Grangier and Yann Dauphin

10:49–11:07  Deep Neural Machine Translation with Linear Associative Unit
Mingxuan Wang, Zhengdong Lu, Jie Zhou and Qun Liu
Monday, July 31st

10:30–11:45 Session 1E: Generation 1

10:30–10:48 Neural AMR: Sequence-to-Sequence Models for Parsing and Generation
Ioannis Konstas, Srinivasan Iyer, Mark Yatskar, Yejin Choi and Luke Zettlemoyer

10:49–11:07 Program Induction by Rationale Generation: Learning to Solve and Explain Algebraic Word Problems
Wang Ling, Dani Yogatama, Chris Dyer and Phil Blunsom

11:08–11:26 Automatically Generating Rhythmic Verse with Neural Networks
Jack Hopkins and Douwe Kiela

11:27–11:45 Creating Training Corpora for NLG Micro-Planners
Claire Gardent, Anastasia Shimorina, Shashi Narayan and Laura Perez-Beltrachini

Monday, July 31st

13:40–14:55 Session 2A: Question Answering 1

13:40–13:58 Gated Self-Matching Networks for Reading Comprehension and Question Answering
Wenhui Wang, Nan Yang, Furu Wei, Baobao Chang and Ming Zhou

13:59–14:17 Generating Natural Answers by Incorporating Copying and Retrieving Mechanisms in Sequence-to-Sequence Learning
Shizhu He, Cao Liu, Kang Liu and Jun Zhao

14:18–14:36 Coarse-to-Fine Question Answering for Long Documents
Eunsol Choi, Daniel Hewlett, Jakob Uszkoreit, Illia Polosukhin, Alexandre Lacoste and Jonathan Berant

14:37–14:55 An End-to-End Model for Question Answering over Knowledge Base with Cross-Attention Combining Global Knowledge
Yanchao Hao, Yuanzhe Zhang, Kang Liu, Shizhu He, Zhanyi Liu, Hua Wu and Jun Zhao
**Monday, July 31st**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session 2B: Vision 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:40–13:58</td>
<td><em>Translating Neuralese</em> Jacob Andreas, Anca Dragan and Dan Klein</td>
</tr>
<tr>
<td>13:59–14:17</td>
<td><em>Obtaining referential word meanings from visual and distributional information: Experiments on object naming</em> Sina Zarrieß and David Schlangen</td>
</tr>
<tr>
<td>14:18–14:36</td>
<td><em>FOIL it! Find One mismatch between Image and Language caption</em> Ravi Shekhar, Sandro Pezzelle, Yauhen Klimovich, Aurélie Herbelot, Moin Nabi, Enver Sangineto and Raffaella Bernardi</td>
</tr>
</tbody>
</table>

**Monday, July 31st**

<table>
<thead>
<tr>
<th>Time</th>
<th>Session 2C: Syntax 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:18–14:36</td>
<td><em>Aggregating and Predicting Sequence Labels from Crowd Annotations</em> An Thanh Nguyen, Byron Wallace, Junyi Jessy Li, Ani Nenkova and Matthew Lease</td>
</tr>
</tbody>
</table>
Monday, July 31st

13:40–15:14 Session 2D: Machine Learning 1 (NN)

Chunting Zhou and Graham Neubig

13:59–14:17 Scalable Bayesian Learning of Recurrent Neural Networks for Language Modeling
Zhe Gan, Chunyuan Li, Changyou Chen, Yunchen Pu, Qinliang Su and Lawrence Carin

14:18–14:36 Learning attention for historical text normalization by learning to pronounce
Marcel Bollmann, Joachim Bingel and Anders Søgaard

Danilo Croce, Simone Filice, Giuseppe Castellucci and Roberto Basili

14:56–15:14 Topically Driven Neural Language Model
Jey Han Lau, Timothy Baldwin and Trevor Cohn

Monday, July 31st

13:40–14:55 Session 2E: Sentiment 1 (NN)

13:40–13:58 Handling Cold-Start Problem in Review Spam Detection by Jointly Embedding Texts and Behaviors
Xuepeng Wang, Kang Liu and Jun Zhao

13:59–14:17 Learning Cognitive Features from Gaze Data for Sentiment and Sarcasm Classification using Convolutional Neural Network
Abhijit Mishra, Kuntal Dey and Pushpak Bhattacharyya

14:18–14:36 An Unsupervised Neural Attention Model for Aspect Extraction
Ruidan He, Wee Sun Lee, Hwee Tou Ng and Daniel Dahlmeier

14:37–14:55 Other Topics You May Also Agree or Disagree: Modeling Inter-Topic Preferences using Tweets and Matrix Factorization
Akira Sasaki, Kazuaki Hanawa, Naoaki Okazaki and Kentaro Inui
Monday, July 31st

15:45–16:41  Session 3A: Information Extraction 2 / Biomedical 1

15:45–16:03  Automatically Labeled Data Generation for Large Scale Event Extraction
Yubo Chen, Shulin Liu, Xiang Zhang, Kang Liu and Jun Zhao

16:04–16:22  Time Expression Analysis and Recognition Using Syntactic Token Types and General Heuristic Rules
Xiaoshi Zhong, Aixin Sun and Erik Cambria

16:23–16:41  Learning with Noise: Enhance Distantly Supervised Relation Extraction with Dynamic Transition Matrix
Bingfeng Luo, Yansong Feng, Zheng Wang, Zhanxing Zhu, Songfang Huang, Rui Yan and Dongyan Zhao

Monday, July 31st

15:45–17:00  Session 3B: Semantics 2 (NN)

15:45–16:03  A Syntactic Neural Model for General-Purpose Code Generation
Pengcheng Yin and Graham Neubig

16:04–16:22  Learning bilingual word embeddings with (almost) no bilingual data
Mikel Artetxe, Gorka Labaka and Eneko Agirre

16:23–16:41  Abstract Meaning Representation Parsing using LSTM Recurrent Neural Networks
William Foland and James H. Martin

16:42–17:00  Deep Semantic Role Labeling: What Works and What’s Next
Luheng He, Kenton Lee, Mike Lewis and Luke Zettlemoyer
Monday, July 31st

15:45–17:00 Session 3C: Speech 1 / Dialogue 1

15:45–16:03 Towards End-to-End Reinforcement Learning of Dialogue Agents for Information Access
Bhuwan Dhingra, Lihong Li, Xiujun Li, Jianfeng Gao, Yun-Nung Chen, Faisal Ahmed and Li Deng

16:04–16:22 Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-Based Chatbots
Yu Wu, Wei Wu, Chen Xing, Ming Zhou and Zhoujun Li

16:23–16:41 Learning Word-Like Units from Joint Audio-Visual Analysis
David Harwath and James Glass

16:42–17:00 Joint CTC/attention decoding for end-to-end speech recognition
Takaaki Hori, Shinji Watanabe and John Hershey

Monday, July 31st

15:45–16:22 Session 3D: Multilingual 1

15:45–16:03 Found in Translation: Reconstructing Phylogenetic Language Trees from Translations
Ella Rabinovich, Noam Ordan and Shuly Wintner

16:04–16:22 Predicting Native Language from Gaze
Yevgeni Berzak, Chie Nakamura, Suzanne Flynn and Boris Katz
Monday, July 31st

15:45–16:03  **Session 3E: Phonology 1**

15:45–16:03  *MORSE: Semantic-ally Drive-n MORpheme SEgment-er*
Tarek Sakakini, Suma Bhat and Pramod Viswanath

Tuesday, August 1st

10:30–11:45  **Session 4A: Information Extraction 3 (NN)**

10:30–10:48  *Deep Pyramid Convolutional Neural Networks for Text Categorization*
Rie Johnson and Tong Zhang

10:49–11:07  *Improved Neural Relation Detection for Knowledge Base Question Answering*
Mo Yu, Wenpeng Yin, Kazi Saidul Hasan, Cicero dos Santos, Bing Xiang and Bowen Zhou

11:08–11:26  *Deep Keyphrase Generation*
Rui Meng, Sanqiang Zhao, Shuguang Han, Daqing He, Peter Brusilovsky and Yu Chi

11:27–11:45  *Attention-over-Attention Neural Networks for Reading Comprehension*
Yiming Cui, Zhipeng Chen, Si Wei, Shijin Wang, Ting Liu and Guoping Hu
Tuesday, August 1st

10:30–11:26  Session 4B: Cognitive Modelling 1 / Vision 2

10:30–10:48  Alignment at Work: Using Language to Distinguish the Internalization and Self-Regulation Components of Cultural Fit in Organizations
Gabriel Doyle, Amir Goldberg, Sameer Srivastava and Michael Frank

10:49–11:07  Representations of language in a model of visually grounded speech signal
Grzegorz Chrupała, Lieke Gelderloos and Afra Alishahi

11:08–11:26  Spectral Analysis of Information Density in Dialogue Predicts Collaborative Task Performance
Yang Xu and David Reitter

Tuesday, August 1st

10:30–12:04  Session 4C: Dialogue 2

10:30–10:48  Affect-LM: A Neural Language Model for Customizable Affective Text Generation
Sayan Ghosh, Mathieu Chollet, Eugene Laksana, Louis-Philippe Morency and Stefan Scherer

10:49–11:07  Domain Attention with an Ensemble of Experts
Young-Bum Kim, Karl Stratos and Dongchan Kim

Tiancheng Zhao, Ran Zhao and Maxine Eskenazi

11:27–11:45  Hybrid Code Networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning
Jason D Williams, Kavosh Asadi and Geoffrey Zweig

11:46–12:04  Generating Contrastive Referring Expressions
Martin Villalba, Christoph Teichmann and Alexander Koller
Tuesday, August 1st

10:30–11:07 Session 4D: Machine Translation 2

10:30–10:48 Modeling Source Syntax for Neural Machine Translation
Junhui Li, Deyi Xiong, Zhaopeng Tu, Muhua Zhu, Min Zhang and Guodong Zhou

10:49–11:07 Sequence-to-Dependency Neural Machine Translation
Shuangzhi Wu, Dongdong Zhang, Nan Yang, Mu Li and Ming Zhou

Tuesday, August 1st

10:30–11:45 Session 4E: Social Media 1

10:30–10:48 Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning
Jing Ma, Wei Gao and Kam-Fai Wong

10:49–11:07 EmoNet: Fine-Grained Emotion Detection with Gated Recurrent Neural Networks
Muhammad Abdul-Mageed and Lyle Ungar

11:08–11:26 Beyond Binary Labels: Political Ideology Prediction of Twitter Users
Daniel Preoțiuc-Pietro, Ye Liu, Daniel Hopkins and Lyle Ungar

11:27–11:45 Leveraging Behavioral and Social Information for Weakly Supervised Collective Classification of Political Discourse on Twitter
Kristen Johnson, Di Jin and Dan Goldwasser
Tuesday, August 1st

13:49–14:39 Session 5A: Multidisciplinary 1

13:49–14:07 A Nested Attention Neural Hybrid Model for Grammatical Error Correction
Jianshu Ji, Qinlong Wang, Kristina Toutanova, Yongen Gong, Steven Truong and Jianfeng Gao

14:08–14:26 TextFlow: A Text Similarity Measure based on Continuous Sequences
Yassine Mrabet, Halil Kilicoglu and Dina Demner-Fushman

14:27–14:39 Friendships, Rivalries, and Trysts: Characterizing Relations between Ideas in Texts
Chenhao Tan, Dallas Card and Noah A. Smith

Tuesday, August 1st

13:30–14:26 Session 5B: Language and Resources 1

13:30–13:48 Polish evaluation dataset for compositional distributional semantics models
Alina Wróblewska and Katarzyna Krasnowska-Kieraś

13:49–14:07 Automatic Annotation and Evaluation of Error Types for Grammatical Error Correction
Christopher Bryant, Mariano Felice and Ted Briscoe

14:08–14:26 Evaluation Metrics for Machine Reading Comprehension: Prerequisite Skills and Readability
Saku Sugawara, Yusuke Kido, Hikaru Yokono and Akiko Aizawa
Tuesday, August 1st

13:30–14:26  Session 5C: Syntax 2 (NN)

13:30–13:48  *A Minimal Span-Based Neural Constituency Parser*
Mitchell Stern, Jacob Andreas and Dan Klein

13:49–14:07  *Semantic Dependency Parsing via Book Embedding*
Weiwei Sun, Junjie Cao and Xiaojun Wan

14:08–14:26  *Neural Word Segmentation with Rich Pretraining*
Jie Yang, Yue Zhang and Fei Dong

Tuesday, August 1st

13:30–14:07  Session 5D: Machine Translation 3 (NN)

Yusuke Oda, Philip Arthur, Graham Neubig, Koichiro Yoshino and Satoshi Nakamura

Yonatan Belinkov, Nadir Durrani, Fahim Dalvi, Hassan Sajjad and James Glass
Tuesday, August 1st

13:30–14:07 Session 5E: Sentiment 2

13:30–13:48 Context-Dependent Sentiment Analysis in User-Generated Videos
Soujanya Poria, Erik Cambria, Devamanyu Hazarika, Navonil Majumder, Amir Zadeh and Louis-Philippe Morency

13:49–14:07 A Multidimensional Lexicon for Interpersonal Stancetaking
Umashanthi Pavalanathan, Jim Fitzpatrick, Scott Kiesling and Jacob Eisenstein

Tuesday, August 1st

15:25–16:21 Session 6A: Information Extraction 4

Jeffrey Lund, Connor Cook, Kevin Seppi and Jordan Boyd-Graber

15:44–16:02 Apples to Apples: Learning Semantics of Common Entities Through a Novel Comprehension Task
Omid Bakhshandeh and James Allen

16:03–16:21 Going out on a limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees
Arzoo Katiyar and Claire Cardie
Tuesday, August 1st

15:25–16:40 Session 6B: Semantics 2 (NN)

15:25–15:43 *Naturalizing a Programming Language via Interactive Learning*
Sida I. Wang, Samuel Ginn, Percy Liang and Christopher D. Manning

15:44–16:02 *Semantic Word Clusters Using Signed Spectral Clustering*
Joao Sedoc, Jean Gallier, Dean Foster and Lyle Ungar

16:03–16:21 *An Interpretable Knowledge Transfer Model for Knowledge Base Completion*
Qizhe Xie, Xuezhe Ma, Zihang Dai and Eduard Hovy

16:22–16:40 *Learning a Neural Semantic Parser from User Feedback*

Tuesday, August 1st

15:25–17:00 Session 6C: Discourse 2 / Dialogue 3

15:25–15:43 *Joint Modeling of Content and Discourse Relations in Dialogues*
Kechen Qin, Lu Wang and Joseph Kim

15:44–16:02 *Argument Mining with Structured SVMs and RNNs*
Vlad Niculae, Joonsuk Park and Claire Cardie

16:03–16:21 *Neural Discourse Structure for Text Categorization*
Yangfeng Ji and Noah A. Smith

16:22–16:40 *Adversarial Connective-exploiting Networks for Implicit Discourse Relation Classification*
Lianhui Qin, Zhisong Zhang, Hai Zhao, Zhiting Hu and Eric Xing

16:41–17:00 *Don’t understand a measure? Learn it: Structured Prediction for Coreference Resolution optimizing its measures*
Iryna Haponchyk and Alessandro Moschitti
Tuesday, August 1st

15:25–16:21 Session 6D: Machine Learning 2

15:25–15:43 Bayesian Modeling of Lexical Resources for Low-Resource Settings
Nicholas Andrews, Mark Dredze, Benjamin Van Durme and Jason Eisner

15:44–16:02 Semi-Supervised QA with Generative Domain-Adaptive Nets
Zhilin Yang, Junjie Hu, Ruslan Salakhutdinov and William Cohen

16:03–16:21 From Language to Programs: Bridging Reinforcement Learning and Maximum Marginal Likelihood
Kelvin Guu, Panupong Pasupat, Evan Liu and Percy Liang

Tuesday, August 1st

15:25–17:00 Session 6E: Summarization 1

15:25–15:43 Diversity driven attention model for query-based abstractive summarization
Preksha Nema, Mitesh M. Khapra, Anirban Laha and Balaraman Ravindran

15:44–16:02 Get To The Point: Summarization with Pointer-Generator Networks
Abigail See, Peter J. Liu and Christopher D. Manning

16:03–16:21 Supervised Learning of Automatic Pyramid for Optimization-Based Multi-Document Summarization
Maxime Peyrard and Judith Eckle-Kohler

16:22–16:40 Selective Encoding for Abstractive Sentence Summarization
Qingyu Zhou, Nan Yang, Furu Wei and Ming Zhou

16:41–17:00 PositionRank: An Unsupervised Approach to Keyphrase Extraction from Scholarly Documents
Corina Florescu and Cornelia Caragea
Wednesday, August 2nd

10:40–11:36  Session 7A: Outstanding Papers 1

10:40–10:58  Towards an Automatic Turing Test: Learning to Evaluate Dialogue Responses
Ryan Lowe, Michael Noseworthy, Iulian Vlad Serban, Nicolas Angelard-Gontier, Yoshua Bengio and Joelle Pineau

10:59–11:17  A Transition-Based Directed Acyclic Graph Parser for UCCA
Daniel Hershcovich, Omri Abend and Ari Rappoport

11:18–11:36  Abstract Syntax Networks for Code Generation and Semantic Parsing
Maxim Rabinovich, Mitchell Stern and Dan Klein

Wednesday, August 2nd

10:40–11:17  Session 7B: Outstanding Papers 2

10:40–10:58  Visualizing and Understanding Neural Machine Translation
Yanzhuo Ding, Yang Liu, Huanbo Luan and Maosong Sun

10:59–11:17  Detecting annotation noise in automatically labelled data
Ines Rehbein and Josef Ruppenhofer
Wednesday, August 2nd

15:00–16:34  Session 8A: Outstanding Papers 3

15:00–15:18  Abstractive Document Summarization with a Graph-Based Attentional Neural Model
              Jiwei Tan, Xiaojun Wan and Jianguo Xiao

              Ryan Cotterell and Jason Eisner

              Xinchi Chen, Zhan Shi, Xipeng Qiu and Xuanjing Huang

15:57–16:15  Neural Joint Model for Transition-based Chinese Syntactic Analysis
              Shuhei Kurita, Daisuke Kawahara and Sadao Kurohashi

16:16–16:34  Robust Incremental Neural Semantic Graph Parsing
              Jan Buys and Phil Blunsom

Wednesday, August 2nd

15:00–16:34  Session 8B: Outstanding Papers 4

15:00–15:18  Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme
              Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou and Bo Xu

              Mingbin Xu, Hui Jiang and Sedtawut Watchawittayakul

15:38–15:56  Vancouver Welcomes You! Minimalist Location Metonymy Resolution
              Milan Gritta, Mohammad Taher Pilehvar, Nut Limspatham and Nigel Collier

15:57–16:15  Unifying Text, Metadata, and User Network Representations with a Neural Network for Geolocation Prediction
              Yasuhide Miura, Motoki Taniguchi, Tomoki Taniguchi and Tomoko Ohkuma

16:16–16:34  Multi-Task Video Captioning with Video and Entailment Generation
              Ramakanth Pasunuru and Mohit Bansal
Monday, July 31st

18:00–21:30 Session P1: Poster Session 1

*Enriching Complex Networks with Word Embeddings for Detecting Mild Cognitive Impairment from Speech Transcripts*
Leandro Santos, Edilson Anselmo Corrêa Júnior, Osvaldo Oliveira Jr, Diego Amancio, Letícia Mansur and Sandra Aluíso

*Adversarial Adaptation of Synthetic or Stale Data*
Young-Bum Kim, Karl Stratos and Dongchan Kim

*Chat Detection in an Intelligent Assistant: Combining Task-oriented and Non-task-oriented Spoken Dialogue Systems*
Satoshi Akasaki and Nobuhiro Kaji

*A Neural Local Coherence Model*
Dat Tien Nguyen and Shafiq Joty

*Data-Driven Broad-Coverage Grammars for Opinionated Natural Language Generation (ONLG)*
Tomer Cagan, Stefan L. Frank and Reut Tsarfaty

*Learning to Ask: Neural Question Generation for Reading Comprehension*
Xinya Du, Junru Shao and Claire Cardie

*Joint Optimization of User-desired Content in Multi-document Summaries by Learning from User Feedback*
Avinesh PVS and Christian M. Meyer

*Flexible and Creative Chinese Poetry Generation Using Neural Memory*
Jiyuan Zhang, Yang Feng, Dong Wang, Yang Wang, Andrew Abel, Shiyue Zhang and Andi Zhang

*Learning to Generate Market Comments from Stock Prices*
Soichiro Murakami, Akihiko Watanabe, Akira Miyazawa, Keiichi Goshima, Toshihiko Yanase, Hiroya Takamura and Yusuke Miyao
Can Syntax Help? Improving an LSTM-based Sentence Compression Model for New Domains
Liangguo Wang, Jing Jiang, Hai Leong Chieu, Chen Hui Ong, Dandan Song and Lejian Liao

Transductive Non-linear Learning for Chinese Hypernym Prediction
Chengyu Wang, Junchi Yan, Aoying Zhou and Xiaofeng He

A Constituent-Centric Neural Architecture for Reading Comprehension
Pengtao Xie and Eric Xing

Cross-lingual Distillation for Text Classification
Ruochen Xu and Yiming Yang

Understanding and Predicting Empathic Behavior in Counseling Therapy
Verónica Pérez-Rosas, Rada Mihalcea, Kenneth Resnicow, Satinder Singh and Lawrence An

Leveraging Knowledge Bases in LSTMs for Improving Machine Reading
Bishan Yang and Tom Mitchell

Prerequisite Relation Learning for Concepts in MOOCs
Liangming Pan, Chengjiang Li, Juanzi Li and Jie Tang

Unsupervised Text Segmentation Based on Native Language Characteristics
Shervin Malmasi, Mark Dras, Mark Johnson, Lan Du and Magdalena Wolska

Weakly Supervised Cross-Lingual Named Entity Recognition via Effective Annotation and Representation Projection
Jian Ni, Georgiana Dinu and Radu Florian

Context Sensitive Lemmatization Using Two Successive Bidirectional Gated Recurrent Networks
Abhisek Chakrabarty, Onkar Arun Pandit and Utpal Garain

Learning to Create and Reuse Words in Open-Vocabulary Neural Language Modeling
Kazuya Kawakami, Chris Dyer and Phil Blunsom

Bandit Structured Prediction for Neural Sequence-to-Sequence Learning
Julia Kreutzer, Artem Sokolov and Stefan Riezler
Monday, July 31st (continued)

**Prior Knowledge Integration for Neural Machine Translation using Posterior Regularization**
Jiacheng Zhang, Yang Liu, Huanbo Luan, Jingfang Xu and Maosong Sun

**Incorporating Word Reordering Knowledge into Attention-based Neural Machine Translation**
Jinchao Zhang, Mingxuan Wang, Qun Liu and Jie Zhou

**Lexically Constrained Decoding for Sequence Generation Using Grid Beam Search**
Chris Hokamp and Qun Liu

**Combating Human Trafficking with Multimodal Deep Models**
Edmund Tong, Amir Zadeh, Cara Jones and Louis-Philippe Morency

**MalwareTextDB: A Database for Annotated Malware Articles**
Swee Kiat Lim, Aldrian Obaja Muis, Wei Lu and Chen Hui Ong

**A Corpus of Annotated Revisions for Studying Argumentative Writing**
Fan Zhang, Homa B. Hashemi, Rebecca Hwa and Diane Litman

**Automatic Induction of Synsets from a Graph of Synonyms**
Dmitry Ustalov, Alexander Panchenko and Chris Biemann

**Neural Modeling of Multi-Predicate Interactions for Japanese Predicate Argument Structure Analysis**
Hiroki Ouchi, Hiroyuki Shindo and Yuji Matsumoto

**TriviaQA: A Large Scale Distantly Supervised Challenge Dataset for Reading Comprehension**
Mandar Joshi, Eunsol Choi, Daniel Weld and Luke Zettlemoyer

**Learning Semantic Correspondences in Technical Documentation**
Kyle Richardson and Jonas Kuhn

**Bridge Text and Knowledge by Learning Multi-Prototype Entity Mention Embedding**
Yixin Cao, Lifu Huang, Heng Ji, Xu Chen and Juanzi Li

**Interactive Learning of Grounded Verb Semantics towards Human-Robot Communication**
Lanbo She and Joyce Chai
Multimodal Word Distributions
Ben Athiwaratkun and Andrew Wilson

Enhanced LSTM for Natural Language Inference
Qian Chen, Xiaodan Zhu, Zhen-Hua Ling, Si Wei, Hui Jiang and Diana Inkpen

Linguistic analysis of differences in portrayal of movie characters
Anil Ramakrishna, Victor R. Martínez, Nikolaos Malandrakis, Karan Singla and Shrikanth Narayanan

Linguistically Regularized LSTM for Sentiment Classification
Qiao Qian, Minlie Huang, Jinhao Lei and Xiaoyan Zhu

Sarcasm SIGN: Interpreting Sarcasm with Sentiment Based Monolingual Machine Translation
Lotem Peled and Roi Reichart

Active Sentiment Domain Adaptation
Fangzhao Wu, Yongfeng Huang and Jun Yan

Volatility Prediction using Financial Disclosures Sentiments with Word Embedding-based IR Models
Navid Rekabsaz, Mihai Lupu, Artem Baklanov, Alexander Dür, Linda Andersson and Allan Hanbury

CANE: Context-Aware Network Embedding for Relation Modeling
Cunchao Tu, Han Liu, Zhiyuan Liu and Maosong Sun

Universal Dependencies Parsing for Colloquial Singaporean English
Hongmin Wang, Yue Zhang, GuangYong Leonard Chan, Jie Yang and Hai Leong Chieu

Generic Axiomatization of Families of Noncrossing Graphs in Dependency Parsing
Anssi Yli-Jyrä and Carlos Gómez-Rodríguez

Semi-supervised sequence tagging with bidirectional language models
Matthew Peters, Waleed Ammar, Chandra Bhagavatula and Russell Power
Tuesday, August 1st

19:00–22:00  Session P2: Poster Session 2

*Learning Symmetric Collaborative Dialogue Agents with Dynamic Knowledge Graph Embeddings*
He He, Anusha Balakrishnan, Mihail Eric and Percy Liang

*Neural Belief Tracker: Data-Driven Dialogue State Tracking*
Nikola Mrkšić, Diarmuid Ó Séaghdha, Tsung-Hsien Wen, Blaise Thomson and Steve Young

*Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms*
Shulin Liu, Yubo Chen, Kang Liu and Jun Zhao

*Topical Coherence in LDA-based Models through Induced Segmentation*
Hesam Amoualian, Wei Lu, Eric Gaussier, Georgios Balikas, Massih R Amini and Marianne Clausel

*Jointly Extracting Relations with Class Ties via Effective Deep Ranking*
Hai Ye, Wenhan Chao, Zhunchen Luo and Zhoujun Li

*Search-based Neural Structured Learning for Sequential Question Answering*
Mohit Iyyer, Wen-tau Yih and Ming-Wei Chang

*Gated-Attention Readers for Text Comprehension*
Bhuwan Dhingra, Hanxiao Liu, Zhihui Yang, William Cohen and Ruslan Salakhutdinov

*Determining Gains Acquired from Word Embedding Quantitatively Using Discrete Distribution Clustering*
Jianbo Ye, Yanran Li, Zhaohui Wu, James Z. Wang, Wenjie Li and Jia Li

*Towards a Seamless Integration of Word Senses into Downstream NLP Applications*
Mohammad Taher Pilehvar, Jose Camacho-Collados, Roberto Navigli and Nigel Collier

*Reading Wikipedia to Answer Open-Domain Questions*
Danqi Chen, Adam Fisch, Jason Weston and Antoine Bordes
Tuesday, August 1st (continued)

*Learning to Skim Text*
Adams Wei Yu, Hongrae Lee and Quoc Le

*An Algebra for Feature Extraction*
Vivek Srikumar

*Chunk-based Decoder for Neural Machine Translation*
Shonosuke Ishiwatari, Jingtao Yao, Shujie Liu, Mu Li, Ming Zhou, Naoki Yosinaga, Masaru Kituregawa and Weijia Jia

*Doubly-Attentive Decoder for Multi-modal Neural Machine Translation*
Iacer Calixto, Qun Liu and Nick Campbell

*A Teacher-Student Framework for Zero-Resource Neural Machine Translation*
Yun Chen, Yang Liu, Yong Cheng and Victor O.K. Li

*Improved Neural Machine Translation with a Syntax-Aware Encoder and Decoder*
Huadong Chen, Shujian Huang, David Chiang and Jiajun Chen

*Cross-lingual Name Tagging and Linking for 282 Languages*
Xiaoman Pan, Boliang Zhang, Jonathan May, Joel Nothman, Kevin Knight and Heng Ji

*Adversarial Training for Unsupervised Bilingual Lexicon Induction*
Meng Zhang, Yang Liu, Huanbo Luan and Maosong Sun

*Estimating Code-Switching on Twitter with a Novel Generalized Word-Level Language Detection Technique*
Shruti Rijhwani, Royal Sequiera, Monojit Choudhury, Kalika Bali and Chandra Shekhar Maddila

*Using Global Constraints and Reranking to Improve Cognates Detection*
Michael Bloodgood and Benjamin Strauss

*One-Shot Neural Cross-Lingual Transfer for Paradigm Completion*
Katharina Kann, Ryan Cotterell and Hinrich Schütze

*Morphological Inflection Generation with Hard Monotonic Attention*
Roee Aharoni and Yoav Goldberg
Tuesday, August 1st (continued)

*From Characters to Words to in Between: Do We Capture Morphology?*
Clara Vania and Adam Lopez

*Riemannian Optimization for Skip-Gram Negative Sampling*
Alexander Fonarev, Oleksii Grinchuk, Gleb Gusev, Pavel Serdyukov and Ivan Oseledets

*Deep Multitask Learning for Semantic Dependency Parsing*
Hao Peng, Sam Thomson and Noah A. Smith

*Improved Word Representation Learning with Sememes*
Yilin Niu, Ruobing Xie, Zhiyuan Liu and Maosong Sun

*Learning Character-level Compositionality with Visual Features*
Frederick Liu, Han Lu, Chieh Lo and Graham Neubig

*A Progressive Learning Approach to Chinese SRL Using Heterogeneous Data*
Qiaolin Xia, Lei Sha, Baobao Chang and Zhifang Sui

*Revisiting Recurrent Networks for Paraphrastic Sentence Embeddings*
John Wieting and Kevin Gimpel

*Ontology-Aware Token Embeddings for Prepositional Phrase Attachment*
Pradeep Dasigi, Waleed Ammar, Chris Dyer and Eduard Hovy

*Identifying 1950s American Jazz Musicians: Fine-Grained IsA Extraction via Modifier Composition*
Ellie Pavlick and Marius Pasca

*Parsing to 1-Endpoint-Crossing, Pagenumber-2 Graphs*
Junjie Cao, Sheng Huang, Weiwei Sun and Xiaojun Wan

*Semi-supervised Multitask Learning for Sequence Labeling*
Marek Rei

*Semantic Parsing of Pre-university Math Problems*
Takuya Matsuzaki, Takumi Ito, Hidenao Iwane, Hirokazu Anai and Noriko H. Arai