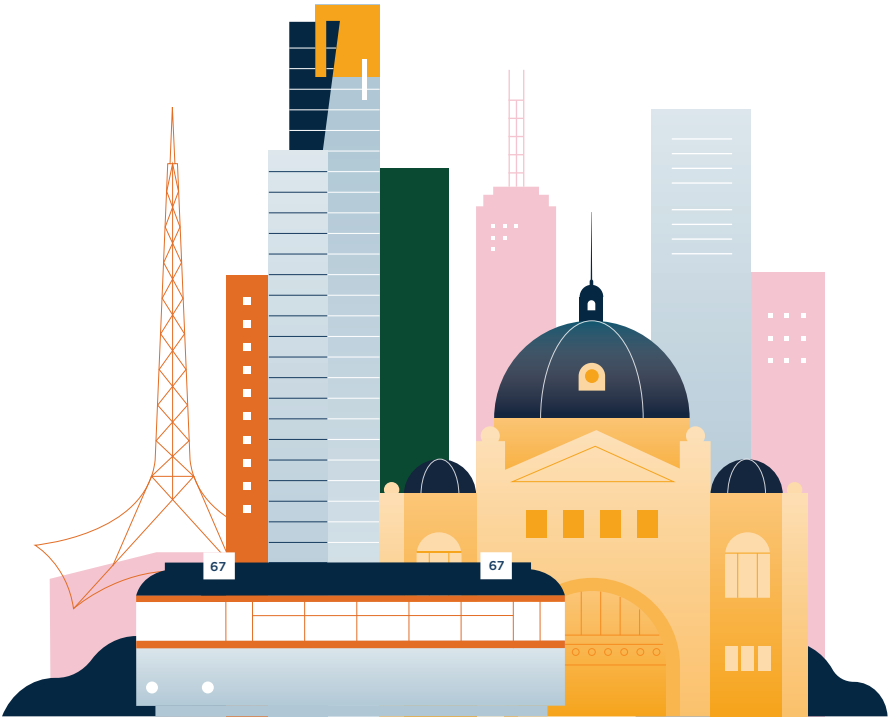
ACL2018

15 – 20 JULY / MELBOURNE AUSTRALIA



CONFERENCE HANDBOOK

56TH ANNUAL MEETING

ASSOCIATION FOR COMPUTATIONAL LINGUISTICS

†: *Start, end and lunch times vary between workshops, the above is a rough guide.*

Cover design by Carrie Chilton
Handbook assembled by Jey Han Lau and Trevor Cohn
Printing by Minuteman Press

Contents

1

Conference Information

Message from the CEO of the Melbourne Convention Bureau

Melbourne Convention Bureau
Collins Square
Tower Two
Level 28, 727 Collins Street
Melbourne VIC 3008

GPO Box 2219T
Melbourne VIC 3001 Australia

Tel +61 3 9002 2222
info@melbournecb.com.au
melbournecb.com.au
ABN 62 072 324 933

On behalf of the Melbourne Convention Bureau, it is my great pleasure to welcome you to Melbourne - officially the world's most liveable city and host of the 56th Annual Meeting of the Association for Computational Linguistics.

Melbourne is renowned as a city that is full of surprises; around every corner, laneway, building and outdoor space, Melbourne is alive with fashion, food, art and culture, and I am excited for you to make these discoveries.

Every Melburnian absolutely loves their city and in the next few days you will find out why. I wish you a very successful and enjoyable visit to Melbourne.

Yours sincerely

KAREN BOLINGER
CHIEF EXECUTIVE OFFICER
MELBOURNE CONVENTION BUREAU

a subsidiary of VISIT VICTORIA

Message from the General Chair

It is an honor to write the initial words of this proceedings as General Chair of the 56th Annual Meeting of the Association for Computational Linguistics! This is only the second time that an ACL conference has been held in Australia — the first time was for the joint COLING/ACL conference in June of 2006 in Sydney, and I was one of its Program Chairs. For ACL 2018 we have tried to maintain the welcoming and intimate spirit and the relaxed and genial character of the much smaller ACL conferences of the past in spite of the ever-growing number of researchers in the field and participants in our conferences.

It is my pleasure here to express gratitude to all those without whom this conference would not exist. My biggest thanks go to the Program Chairs Iryna Gurevych and Yusuke Miyao, as well as to Local Chairs Tim Baldwin, Trevor Cohn and Karin Verspoor. They have done a tremendous job to manage the submission and review process, and the local arrangement details, respectively.

I also want to thank all of the other chairs for their very hard work: Workshops Chairs Brendan O'Connor and Eva Maria Vecchi; Tutorials Chairs Yoav Artzi and Jacob Eisenstein; Demo Chairs Fei Liu and Thamar Solorio; Student Research Workshop Organizers Vered Shwartz, Jeniya Tabassum and Rob Voigt; Faculty Advisors to the Student Research Workshop Marie-Catherine de Marneffe, Wanxiang Che and Malvina Nissim; Publications Chairs Shay Cohen, Kevin Gimpel and Wei Lu; Exhibits Coordinator Karin Verspoor; Student Volunteer Coordinator Karin Vespoor; Conference Handbook Chairs Jey Han Lau and Trevor Cohn; Publicity Chair Sarvnaz Karimi; Local Sponsorship Chair Cecile Paris; Webmaster Andrew MacKinlay; and Priscilla Rasmussen, giver of advice and wisdom to all of us as ACL Business Manager.

I also warmly thank the ACL Executive Committee for its guidance and advice on many important issues and concerns as they arose.

I am also extremely grateful to all the sponsors for their great support to the conference.

Many thanks to the area chairs, the reviewers, the invited speakers, the authors of the various papers, posters and presentations.

And, finally, many many thanks to all the participants who will put the final touches on making ACL 2018 an exciting, stimulating and inspiring event!

ACL 2018 General Chair
Claire Cardie, Cornell University

Message from the Program Committee Co-Chairs

Welcome to the 56th Annual Meeting of the Association for Computational Linguistics 2018 – or ACL 2018 for short.

In September 2017, Program Committee Co-Chairs (PCs) posted the call for nominations of Area Chairs (AC), Reviewers and Invited Speakers. We received 752 responses in total. Overall, out of 388 valid nominations for area chairs, 299 unique persons were suggested; 110 persons were self-nominations. About 70% of the 56 selected area chairs (later expanded to 61 area chairs due to the high number of submissions) were nominated by the community. For the reviewers, we collected 936 valid nominations. At the PhD level, 139 persons were self-nominations and 129 were nominated by others. At the Postdoc/Ass.Prof. level, 160 were self-nominated, 112 nominated by others. At the Prof. level, 221 persons were self-nominated, 175 nominated by others.

We received 138 unique nominations for invited speakers, from which two invited speakers of the conference were selected:

- Carolyn Penstein Rosé, Language Technologies Institute at Carnegie Mellon University, USA
- Anton van den Hengel, Australian Centre for Visual Technologies at University of Adelaide, Australia

Our community is steadily growing: in total, 1621 submissions were received right after the submission deadline: 1045 long, 576 short papers. 13 erroneous submissions were deleted or withdrawn in the preliminary checks by PCs. 25 papers were rejected without review (16 long, 9 short); the reasons are the violation of the ACL 2018 style and dual submission guidelines. 32 papers were withdrawn before the review period started; the main reason was that the papers have been accepted as the short papers at NAACL HLT 2018. In total, 1551 papers went into the reviewing phase: 1021 long, 530 short papers. 1610 reviewers (1473 primary and 137 secondary reviewers) were involved in the reviewing process; each reviewer has reviewed about 3 papers on average. 3 long and 4 short papers were withdrawn during the reviewing period, and finally 1018 long and 526 short papers were considered during the acceptance decision phase.

The assignment of papers to areas and reviewers has been done in multiple rounds. First round: Initial assignments of papers to areas were determined automatically with the help of the authors' input, while PCs went through all submissions and moved papers to other areas, considering COI and the topical fit. PCs assigned one AC as a meta-reviewer to each paper using Toronto Paper Matching System (TPMS) scores. Second round: ACs looked into the papers in their area, and adjusted meta-reviewer assignments. ACs sent a report to PCs if they found any problems. Third round: PCs made the final decision, considering the workload balance, possible COIs and the topical fit. Fourth round: ACs decided which reviewers would review each paper, based on AC's knowledge about the reviewers, TPMS scores, reviewers' bids, and COI.

We have introduced several innovations to the reviewing process. One of them is an argument-based review form. The reviewers were asked to provide arguments for and against the paper. This has been tremendously helpful for ACs and PCs to analyze the reviews and come up with final recommendations. The authors were asked to respond to the con arguments during the rebuttal. In coordination with the NAACL HLT 2018 PCs, we plan to do some analytics on anonymized reviews and rebuttal statements, with the consent of the reviewers and authors. Our purpose is to improve the quality of the review process. The data will be compiled into a unique corpus for NLP, and will be made available to the research community after appropriate anonymization checks, at the earliest in 2 years after ACL 2018. We hope to provide data on *how to review* to younger researchers, and to improve the transparency of the reviewing process in general.

The ACL 2018 conference is super-competitive: We accepted 256 out of 1018 submitted long papers and 125 out of 526 short papers, with an overall acceptance rate of 24.7%. The details of the review process are available at the conference homepage. Criteria of acceptance were mainly:

- strengths/weaknesses raised by reviewers and their significance;
- the result of discussions and author responses;
- contribution to CL as the science of language: whether the paper advances (or contributes to) our understanding of language in any way;
- diversity: we do not want to fill ACL with similar papers like achieving 1% improvement on a well-known task.

We also considered the balance of paper types, topics and contributions and re-considered the acceptance when reviewers reported any problem in preliminary checks (*Appropriateness* to *Handling of Human Participants*).

Continuing the tradition, ACL 2018 will feature 20 papers which were accepted for publication in the Transactions of the Association for Computational Linguistics (TACL). The TACL papers were split into 10 oral presentations and 10 poster presentations.

There are many people to thank for who have worked diligently to make ACL 2018 possible. All names are listed in the Program Committee section of the Front Matter.

Since the conference size continues to grow and the organizational complexity increases, we have introduced the role of Program Committee Co-Chair Assistants. In total, 5 senior researchers have supported the PCs during most intensive work phases to handle the communication in a timely manner, draft various documents and effectively prepare decisions.

Thanks to our area chairs for their hard work on recruiting reviewers, managing reviews, leading discussions, and making recommendations.

This program certainly would not be possible without the help of the 1610 reviewers. In particular, 192 reviewers from this list were recognized by the area chairs as outstanding reviewers who have turned in exceptionally well-written and constructive reviews and who have actively engaged themselves in the post-rebuttal discussions.

We are also deeply indebted to the best paper selection committee which consists of 22 members. They had to additionally review 6-8 papers according to the best paper criteria on short notice. Their time and effort in recommending the best paper awards is much appreciated.

We also would like to thank many colleagues for generously sharing their experience in organizing prior ACL conferences and for their advice. We are grateful for the guidance and the support of the ACL presidents Joakim Nivre and Marti Hearst, and the ACL board. We also would like to thank the publication co-chairs Shay Cohen, Kevin Gimpel and Wei Lu (Advisory) and the handbook chair Jey Han Lau for putting together the proceedings and the conference handbook; and Rich Gerber from Softconf for always being responsive to our requests. We would like to thank the ACL Business Manager Priscilla Rasmussen for helping us to sort important things out. Finally, this conference could not have happened without the efforts of the general chair, Claire Cardie. We thank her for the leadership and advice, especially when matters got complicated.

We hope you will enjoy ACL 2018 and contribute to the future success of our community!

ACL 2018 Program Committee Co-Chairs
Iryna Gurevych, TU Darmstadt, Germany
Yusuke Miyao, National Institute of Informatics, Japan

Organizing Committee

General Chair
Claire Cardie, Cornell University

Program Chairs
Iryna Gurevych, TU Darmstadt
Yusuke Miyao, National Institute of Informatics

Workshop Chairs
Brendan O'Connor, University of Massachusetts Amherst
Eva Maria Vecchi, University of Cambridge

Tutorial Chairs
Yoav Artzi, Cornell University
Jacob Eisenstein, Georgia Institute of Technology

Demo Chairs
Fei Liu, University of Central Florida
Thamar Solorio, University of Houston

Publication Chairs
Shay Cohen, University of Edinburgh
Kevin Gimpel, Toyota Technological Institute at Chicago
Wei Lu, Singapore University of Technology and Design (Advisory)

Exhibits Coordinator
Karin Verspoor, University of Melbourne

Conference Handbook Chairs
Jey Han Lau, IBM Research
Trevor Cohn, University of Melbourne

Publicity Chair
Sarvnaz Karimi, CSIRO

Local Sponsorship Chair
Cecile Paris, CSIRO

Student Volunteer Coordinator
Karin Verspoor, University of Melbourne

Local Chairs
Tim Baldwin, University of Melbourne
Karin Verspoor, University of Melbourne
Trevor Cohn, University of Melbourne

Student Research Workshop Organisers
Vered Shwartz, Bar-Ilan University
Jeniya Tabassum, Ohio State University
Rob Voigt, Stanford University

Faculty Advisors to the Student Research Workshop
Marie-Catherine de Marneffe, Ohio State University
Wanxiang Che, Harbin Institute of Technology
Malvina Nissim, University of Groningen

Webmaster
Andrew MacKinlay, Culture Amp / University of Melbourne

Program Committee

Program Committee Co-chairs
Iryna Gurevych, TU Darmstadt, Germany
Yusuke Miyao, National Institute of Informatics, Japan

Program Committee Co-Chair Assistants
Yang Gao, TU Darmstadt, Germany
Ivan Habernal, TU Darmstadt, Germany
Sang Phan, National Institute of Informatics, Japan
Steffen Eger, TU Darmstadt, Germany
Christian Meyer, TU Darmstadt, Germany

Area Chairs
(Senior Area Chairs are indicated in boldface.)

Dialogue and Interactive Systems
- **Asli Celikyilmaz**
- Verena Rieser
- Milica Gasic
- Jason Williams

Discourse and Pragmatics
- Manfred Stede
- **Ani Nenkova**

Document Analysis
- **Hang Li**
- Yiqun Liu
- Eugene Agichtein

Generation
- Ioannis Konstas
- **Claire Gardent**

Information Extraction and Text Mining
- Feiyu Xu
- Kevin Cohen
- Zhiyuan Liu
- **Ralph Grishman**
- Yi Yang
- Nazli Goharian

Linguistic Theories, Cognitive Modeling and Psycholinguistics
- **Shuly Wintner**
- **Tim O'Donnell**

Machine Learning
- Andre Martins
- Ariadna Quattoni
- **Jun Suzuki**

Machine Translation
- Yang Liu (Tsinghua University)
- **Matt Post**
- Lucia Specia
- Dekai Wu

Multidisciplinary and Area Chair COI
- **Yoav Goldberg**
- **Anders Søgaard**
- **Mirella Lapata**

Multilinguality

Bernardo Magnini
Tristan Miller

Phonology, Morphology, and Word Segmentation
Graham Neubig
Hai Zhao

Question Answering
Lluís Màrquez
Teruko Mitamura
Zornitsa Kozareva
Richard Socher

Resources and Evaluation
Gerard de Melo
Sara Tonelli
Karën Fort

Sentence-level Semantics
Luke Zettlemoyer
Ellie Pavlick
Jacob Uszkoreit

Sentiment Analysis and Argument Mining
Smaranda Muresan
Benno Stein
Yulan He

Social Media
David Jurgens
Jing Jiang

Summarization
Kathleen McKeown
Xiaodan Zhu

Tagging, Chunking, Syntax, and Parsing
Liang Huang
Weiwei Sun
Željko Agić
Yue Zhang

Textual Inference and Other Areas of Semantics
Michael Roth
Fabio Massimo Zanzotto

Vision, Robotics, Multimodal, Grounding and Speech
Yoav Artzi
Shinji Watanabe
Timothy Hospedales

Word-level Semantics
Ekaterina Shutova
Roberto Navigli

Best Paper Selection Committee
Timothy Baldwin
Pushpak Bhattacharyya
Phil Blunsom
Johan Bos
Jordan Boyd-Graber
Trevor Cohn
Vera Demberg
Kevin Duh

Katrin Erk
Mark Johnson
Yang Liu (Tsinghua University)
Yuji Matsumoto
Jong Park
Ellie Pavlick
Simone Paolo Ponzetto
Sebastian Riedel
Carolyn Penstein Rosé
Noah A. Smith
Anders Søgaard
Ivan Titov
Benjamin Van Durme
Ming Zhou

Meal Info

The following meals are provided as part of your registration fee:

- Mid-morning breaks include coffee, tea, water, and snacks in the Melbourne Room and level 2 foyer
- A light lunch of sandwiches and a muffin is provided during each of the midday poster sessions on Monday–Wednesday, in the Melbourne Room and level 2 foyer
- Mid-afternoon breaks include coffee, tea, water, and snacks in the Melbourne Room and level 2 foyer

Light refreshments will be provided for students on Monday at the recruitment event, and there is also the Welcome Reception and the Social Event, but you are otherwise on your own.

2

Tutorials: Sunday, July 15

Overview

07:30–18:00	**Registration**	*Level 2 Foyer*
09:00–12:30	**Morning Tutorials**	
	100 Things You Always Wanted to Know about Semantics & Pragmatics But Were Afraid to Ask *Emily M. Bender*	*216*
	Neural Approaches to Conversational AI *Jianfeng Gao, Michel Galley, and Lihong Li*	*219*
	Variational Inference and Deep Generative Models *Wilker Aziz and Philip Schulz*	*217*
	Connecting Language and Vision to Actions *Peter Anderson, Abhishek Das, and Qi Wu*	*218*
10:30–11:00	**Coffee Break**	*Level 2 Foyer*
12:30–13:30	**Lunch Break**	
13:30–17:00	**Afternoon Tutorials**	
	Beyond Multiword Expressions: Processing Idioms and Metaphors *Valia Kordoni*	*216*
	Neural Semantic Parsing *Luke Zettlemoyer, Matt Gardner, Pradeep Dasigi, Srinivasan Iyer, and Alane Suhr*	*217*
	Deep Reinforcement Learning for NLP *William Yang Wang, Jiwei Li, and Xiaodong He*	*219*
	Multi-lingual Entity Discovery and Linking *Avirup Sil, Heng Ji, Dan Roth, and Silviu-Petru Cucerzan*	*218*
15:00–15:30	**Coffee Break**	*Level 2 Foyer*
18:00–20:00	**Welcome Reception**	*Melbourne Room 1*

Message from the Tutorial Co-Chairs

This volume contains the abstracts of the ACL 2018 tutorials. Tutorials were selected from 49 submissions to a joint call, which was coordinated with NAACL, COLING, and EMNLP. From these submissions, eight half-day tutorials were selected for ACL, on the criteria of quality, relevance, interest, and balance. We thank Mausam for coordinating this process across all four conferences, and we wish to acknowledge support from the publications chairs Kevin Gimpel, Shay Cohen, and Wei Lu (ACL publications chairs) and Jey Han Lau and Trevor Cohn (ACL handbook chairs), as well as Stephanie Lukin (NAACL publications co-chair). Most importantly, we thank the tutorial presenters for their contributions, which we hope that you will enjoy.

ACL 2018 Tutorial Co-Chairs
Yoav Artzi, Cornell University Jacob Eisenstein, Georgia Institute of Technology

Tutorial 1

100 Things You Always Wanted to Know about Semantics & Pragmatics But Were Afraid to Ask

Emily M. Bender

Sunday, July 15, 2018, 9:00–12:30

216

Meaning is a fundamental concept in Natural Language Processing (NLP), given its aim to build systems that mean what they say to you, and understand what you say to them. In order for NLP to scale beyond partial, task-specific solutions, it must be informed by what is known about how humans use language to express and understand communicative intents. The purpose of this tutorial is to present a selection of useful information about semantics and pragmatics, as understood in linguistics, in a way that's accessible to and useful for NLP practitioners with minimal (or even no) prior training in linguistics. The tutorial content is based on a manuscript in progress I am co-authoring with Prof. Alex Lascarides of the University of Edinburgh.

Emily M. Bender is a Professor in the Department of Linguistics and Adjunct Professor in the Paul G. Allen School of Computer Science & Engineering at the University of Washington. She is also the past chair (2016–2017) of NAACL. Her research interests lie in multilingual grammar engineering, computational semantics, and the incorporation of linguistic knowledge in natural language processing. She is the primary developer of the Grammar Matrix grammar customization system, which is developed in the context of the DELPH-IN Consortium (Deep Linguistic Processing with HPSG Initiative). More generally, she is interested in the intersection of linguistics and computational linguistics, from both directions: bringing computational methodologies to linguistic science and linguistic science to natural language processing.

Tutorial 2

Neural Approaches to Conversational AI

Jianfeng Gao, Michel Galley, and Lihong Li

Sunday, July 15, 2018, 9:00–12:30

219

This tutorial surveys neural approaches to conversational AI that were developed in the last few years. We group conversational systems into three categories: (1) question answering agents, (2) task-oriented dialogue agents, and (3) social bots. For each category, we present a review of state-of-the-art neural approaches, draw the connection between neural approaches and traditional symbolic approaches, and discuss the progress we have made and challenges we are facing, using specific systems and models as case studies.

Jianfeng Gao is Partner Research Manager at Microsoft AI and Research, Redmond. He leads the development of AI systems for machine reading comprehension, question answering, chitchat bots, task-oriented dialogue, and business applications. From 2014 to 2017, he was Partner Research Manager and Principal Researcher at Deep Learning Technology Center at Microsoft Research, Redmond, where he was leading the research on deep learning for text and image processing. From 2006 to 2014, he was Researcher, Senior Researcher, and Principal Researcher at Natural Language Processing Group at Microsoft Research, Redmond, where he worked on the Bing search engine, improving its core relevance engine and query spelling, understanding and reformulation engines, MS ads relevance and prediction, and statistical machine translation. From 2005 to 2006, he was a Research Lead in Natural Interactive Services Division at Microsoft, where he worked on Project X, an effort of developing natural user interface for Windows. From 2000 to 2005, he was Research Lead in Natural Language Computing Group at Microsoft Research Asia, where he and his colleagues developed the first Chinese speech recognition system released with Microsoft Office, the Chinese/Japanese Input Method Editors (IME) which were the leading products in the market, and the natural language platform for Microsoft Windows.

Michel Galley is a Senior Researcher at Microsoft Research. His research interests are in the areas of natural language processing and machine learning, with a particular focus on conversational AI, text generation, and machine translation. From 2007 to 2010, he was a Postdoctoral Scholar then Research Associate in the Computer Science department at Stanford University, working primarily on Machine Translation. In 2007, he obtained his Ph.D. from the Computer Science department at Columbia University, with research focusing on summarization, discourse, and dialogue. From 2003 to 2005, he did several internships at USC/ISI and Language Weaver on machine translation, which included work that won several NIST MT competitions. From 2000-2001, he did an 8-month internship and undergraduate thesis work in the Spoken Dialog Systems group at Bell Labs, working on generation for dialogue systems.

Lihong Li is a Research Scientist at Google Inc. He obtained a PhD degree in Computer Science from Rutgers University, specializing in reinforcement learning theory and algorithms. After that, he has held Researcher, Senior Researcher, and Principal Researcher positions in Yahoo! Research (2009-2012) and Microsoft Research (2012-2017), before joining Google. His main research interests are reinforcement learning (in both Markov decision processes and contextual bandits) and other related problems in AI (including active leaning, online learning and large-scale machine learning). His work has found applications in recommendation, advertising, Web search and conversation systems, and has won best paper awards at ICML, AISTATS and WSDM. In recent years, he served as area chairs or senior program

committee members at AAAI, AISTATS, ICML, IJCAI and NIPS. More information can be found on his homepage: `http://lihongli.github.io`.

Tutorial 3

Variational Inference and Deep Generative Models

Wilker Aziz and Philip Schulz

Sunday, July 15, 2018, 9:00–12:30

217

NLP has seen a surge in neural network models in recent years. These models provide state-of-the-art performance on many supervised tasks. Unsupervised and semi-supervised learning has only been addressed scarcely, however. Deep generative models (DGMs) make it possible to integrate neural networks with probabilistic graphical models. Using DGMs one can easily design latent variable models that account for missing observations and thereby enable unsupervised and semi-supervised learning with neural networks. The method of choice for training these models is variational inference.

This tutorial offers a general introduction to variational inference followed by a thorough and example-driven discussion of how to use variational methods for training DGMs. It provides both the mathematical background necessary for deriving the learning algorithms as well as practical implementation guidelines. Importantly, the tutorial will cover models with continuous and discrete variables.

We provide practical coding exercises implemented in IPython notebooks as well as short notes on the more intricate mathematical details that the audience can use as a reference after the tutorial. We expect that with these additional materials the tutorial will have a long-lasting impact on the community.

Wilker Aziz is a research associate at the University of Amsterdam (UvA) working on natural language processing problems such as machine translation, textual entailment, and paraphrasing. His research interests include statistical learning, probabilistic models, and methods for approximate inference. Before joining UvA, Wilker worked on exact sampling and optimisation for statistical machine translation at the University of Sheffield (UK) and at the University of Wolverhampton (UK) where he obtained his PhD. Wilker's background is in Computer Engineering which he studied at the Engineering School of the University of São Paulo (Brazil).

Philip Schulz is an applied scientist at Amazon Research. Before joining Amazon, Philip did his PhD at the University of Amsterdam. During the last months of his PhD trajectory, he visited the University of Melbourne. Philip's background is in Linguistics which he studied at the University of Tübingen and UCL in London. These days, his research interests revolve around statistical learning. He has worked on Bayesian graphical models for machine translation. More recently he has extended this line of work towards deep generative models. More broadly, Philip is interested in probabilistic modeling, approximate inference methods and statistical theory.

Tutorial 4

Connecting Language and Vision to Actions

Peter Anderson, Abhishek Das, and Qi Wu

Sunday, July 15, 2018, 9:00–12:30

218

A long-term goal of AI research is to build intelligent agents that can see the rich visual environment around us, communicate this understanding in natural language to humans and other agents, and act in a physical or embodied environment. To this end, recent advances at the intersection of language and vision have made incredible progress – from being able to generate natural language descriptions of images/videos, to answering questions about them, to even holding free-form conversations about visual content! However, while these agents can passively describe images or answer (a sequence of) questions about them, they cannot act in the world (what if I cannot answer a question from my current view, or I am asked to move or manipulate something?). Thus, the challenge now is to extend this progress in language and vision to embodied agents that take actions and actively interact with their visual environments. To reduce the entry barrier for new researchers, this tutorial will provide an overview of the growing number of multimodal tasks and datasets that combine textual and visual understanding. We will comprehensively review existing state-of-the-art approaches to selected tasks such as image captioning, visual question answering (VQA) and visual dialog, presenting the key architectural building blocks (such as co-attention) and novel algorithms (such as cooperative/adversarial games) used to train models for these tasks. We will then discuss some of the current and upcoming challenges of combining language, vision and actions, and introduce some recently-released interactive 3D simulation environments designed for this purpose.

Peter Anderson is a final year PhD candidate in Computer Science at the Australian National University, supervised by Dr Stephen Gould, and a researcher within the Australian Centre for Robotic Vision (ACRV). His PhD focuses on deep learning for visual understanding in natural language. He was an integral member of the team that won first place in the 2017 Visual Question Answering (VQA) challenge at CVPR, and he has made several contributions in image captioning, including achieving first place on the COCO leaderboard in July 2017. He has published at CVPR, ECCV, EMNLP and ICRA, and spent time at numerous universities and research labs including Adelaide University, Macquarie University, and Microsoft Research. His research is currently focused on vision-and-language understanding in complex 3D environments.

Abhishek Das is a Computer Science PhD student at Georgia Institute of Technology, advised by Dhruv Batra, and working closely with Devi Parikh. He is interested in deep learning and its applications in building agents that can see (computer vision), think (reasoning and interpretability), talk (language modeling) and act (reinforcement learning). He is a recipient of an Adobe Research Fellowship and a Snap Research Fellowship. He has published at CVPR, ICCV, EMNLP, HCOMP and CVIU, co-organized the NIPS 2017 workshop on Visually-Grounded Interaction and Language, and has held visiting positions at Virginia Tech, Queensland Brain Institute and Facebook AI Research. He graduated from Indian Institute of Technology Roorkee in 2015 with a Bachelor's in Electrical Engineering.

Qi Wu is a research fellow in the Australia Centre for Robotic Vision (ACRV) in the University of Adelaide. Before that, he was a postdoc researcher in the Australia Centre for Visual Technologies (ACVT) in the University of Adelaide. He obtained his PhD degree in 2015 and MSc degree in 2011, in Computer Science from University of Bath, United Kingdom. His research interests are mainly in Computer Vision

and Machine Learning. Currently, he is working on the vision to language problem and he is especially an expert in the area of Image Captioning and Visual Question Answering (VQA). His attributes-based image captioning model got first place on the COCO Image Captioning Challenge Leader Board in the October of 2015. He has published several papers in prestigious conferences and journals, such as TPAMI, CVPR, ICCV, ECCV, IJCAI and AAAI.

Tutorial 5

Beyond Multiword Expressions: Processing Idioms and Metaphors

Valia Kordoni

Sunday, July 15, 2018, 13:30–17:00

216

Idioms and metaphors are characteristic to all areas of human activity and to all types of discourse. Their processing is a rapidly growing area in NLP, since they have become a big challenge for NLP systems. Their omnipresence in language has been established in a number of corpus studies and the role they play in human reasoning has also been confirmed in psychological experiments. This makes idioms and metaphors an important research area for computational and cognitive linguistics, and their automatic identification and interpretation indispensable for any semantics-oriented NLP application.

This tutorial aims to provide attendees with a clear notion of the linguistic characteristics of idioms and metaphors, computational models of idioms and metaphors using state-of-the-art NLP techniques, their relevance for the intersection of deep learning and natural language processing, what methods and resources are available to support their use, and what more could be done in the future. Our target audience are researchers and practitioners in machine learning, parsing (syntactic and semantic) and language technology, not necessarily experts in idioms and metaphors, who are interested in tasks that involve or could benefit from considering idioms and metaphors as a pervasive phenomenon in human language and communication.

Valia Kordoni is a professor at Humboldt University Berlin (Deputy Chair for the subject area "English Linguistics"). She is a leader in EU-funded research in Machine Translation, Computational Semantics, and Machine Learning. She has organized conferences and workshops dedicated to research on MWEs, recently including the EACL 2014 10th Workshop on Multiword Expressions (MWE 2014) in Gothenburg, Sweden, the NAACL 2015 11th Workshop on Multiword Expressions in Denver, Colorado, and the ACL 2016 12th Workshop on Multiword Expressions in Berlin, Germany, among others. She has been the Local Chair of ACL 2016 - The 54th Annual Meeting of the Association for Computational Linguistics which took place at the Humboldt University Berlin in August 2016. Recent activities of hers include a tutorial on Robust Automated Natural Language Processing with Multiword Expressions and Collocations in ACL 2013, as well as a tutorial on Beyond Words: Deep Learning for Multiword Expressions and Collocations in ACL 2017. She is the author of Multiword Expressions - From Linguistic Analysis to Language Technology Applications (to appear, Springer).

Tutorial 6

Neural Semantic Parsing

Luke Zettlemoyer, Matt Gardner, Pradeep Dasigi, Srinivasan Iyer, and Alane Suhr

Sunday, July 15, 2018, 13:30–17:00

217

Semantic parsing, the study of translating natural language utterances into machine-executable programs, is a well-established research area and has applications in question answering, instruction following, voice assistants, and code generation. In the last two years, the models used for semantic parsing have changed dramatically with the introduction of neural encoder-decoder methods that allow us to rethink many of the previous assumptions underlying semantic parsing. We aim to inform those already interested in semantic parsing research of these new developments in the field, as well as introduce the topic as an exciting research area to those who are unfamiliar with it.

Current approaches for neural semantic parsing share several similarities with neural machine translation, but the key difference between the two fields is that semantic parsing translates natural language into a formal language, while machine translation translates it into a different natural language. The formal language used in semantic parsing allows for constrained decoding, where the model is constrained to only produce outputs that are valid formal statements. We will describe the various approaches researchers have taken to do this. We will also discuss the choice of formal languages used by semantic parsers, and describe why much recent work has chosen to use standard programming languages instead of more linguistically-motivated representations. We will then describe a particularly challenging setting for semantic parsing, where there is additional context or interaction that the parser must take into account when translating natural language to formal language, and give an overview of recent work in this direction. Finally, we will introduce some tools available in AllenNLP for doing semantic parsing research.

Matt Gardner is a research scientist at the Allen Institute for Artificial Intelligence. His research focuses on question answering and semantic parsing. He is the lead maintainer of the AllenNLP toolkit and a host of the NLP Highlights podcast.

Pradeep Dasigi is a PhD student at the Language Technologies Institute in Carnegie Mellon University. His research interest lies in building knowledge-aware language understanding systems, with a recent focus on neural semantic parsing.

Srinivasan Iyer is a graduate student in the Natural Language Processing group at the University of Washington, Seattle. His main research area is context dependent semantic parsing directly from natural language to general purpose programming source code. Other aspects of his research are learning semantic parsers from massive online resources and incorporating user feedback for model improvement.

Alane Suhr is a PhD student in Computer Science at Cornell University. Alane's research interests include developing machine learning methods for understanding natural language grounded in complex environments and interactions. She is a recipient of an NSF Graduate Research Fellowship, the Best Resource Paper award at ACL 2017, and an Outstanding Paper Award at NAACL 2018.

Luke Zettlemoyer is an Associate Professor in the Paul G. Allen School of Computer Science & Engineering at the University of Washington. He has a been doing research in semantic parsing for many years, and recently shifted to studying neural models for this problem. Luke's honors include multiple best paper awards, a PECASE award, and an Allen Distinguished Investigator award.

Tutorial 7

Deep Reinforcement Learning for NLP

William Yang Wang, Jiwei Li, and Xiaodong He

Sunday, July 15, 2018, 13:30–17:00

219

Many Natural Language Processing (NLP) tasks (including generation, language grounding, reasoning, information extraction, coreference resolution, and dialog) can be formulated as deep reinforcement learning (DRL) problems. However, since language is often discrete and the space for all sentences is infinite, there are many challenges for formulating reinforcement learning problems of NLP tasks. In this tutorial, we provide a gentle introduction to the foundation of deep reinforcement learning, as well as some practical DRL solutions in NLP. We describe recent advances in designing deep reinforcement learning for NLP, with a special focus on generation, dialogue, and information extraction. Finally, we discuss why they succeed, and when they may fail, aiming at providing some practical advice about deep reinforcement learning for solving real-world NLP problems.

William Yang Wang is an Assistant Professor at the Department of Computer Science, University of California, Santa Barbara. He received his PhD from School of Computer Science, Carnegie Mellon University. He focuses on information extraction and he is the faculty author of DeepPath—the first deep reinforcement learning system for multi-hop reasoning. He has published more than 50 papers at leading conferences and journals including ACL, EMNLP, NAACL, CVPR, COLING, IJCAI, CIKM, ICWSM, SIGDIAL, IJCNLP, INTERSPEECH, ICASSP, ASRU, SLT, Machine Learning, and Computer Speech & Language, and he has received paper awards and honors from CIKM, ASRU, and EMNLP. Website: `http://www.cs.ucsb.edu/~william/`.

Jiwei Li recently spent three years and received his PhD in Computer Science from Stanford University. His research interests are deep learning and dialogue. He is the most prolific NLP/ML first author during 2012-2016, and the lead author of the first study in deep reinforcement learning for dialogue generation. He is the recipient of a Facebook Fellowship in 2015. Website: `https://web.stanford.edu/~jiweil/`.

Xiaodong He is the Deputy Managing Director of JD AI Research and Head of the Deep learning, NLP and Speech Lab, and a Technical Vice President of JD.com. He is also an Affiliate Professor at the University of Washington (Seattle), serves in doctoral supervisory committees. Before joining JD.com, He was with Microsoft for about 15 years, served as Principal Researcher and Research Manager of the DLTC at Microsoft Research, Redmond. His research interests are mainly in artificial intelligence areas including deep learning, natural language, computer vision, speech, information retrieval, and knowledge representation. He has published more than 100 papers in ACL, EMNLP, NAACL, CVPR, SIGIR, WWW, CIKM, NIPS, ICLR, ICASSP, Proc. IEEE, IEEE TASLP, IEEE SPM, and other venues. He received several awards including the Outstanding Paper Award at ACL 2015. Website: `http://air.jd.com/people2.html`.

Tutorial 8

Multi-lingual Entity Discovery and Linking

Avirup Sil, Heng Ji, Dan Roth, and Silviu-Petru Cucerzan

Sunday, July 15, 2018, 13:30–17:00

218

The primary goals of this tutorial are to review the framework of cross-lingual EL and motivate it as a broad paradigm for the Information Extraction task. We will start by discussing the traditional EL techniques and metrics and address questions relevant to the adequacy of these to across domains and languages. We will then present more recent approaches such as Neural EL, discuss the basic building blocks of a state-of-the-art neural EL system and analyze some of the current results on English EL. We will then proceed to Cross-lingual EL and discuss methods that work across languages. In particular, we will discuss and compare multiple methods that make use of multi-lingual word embeddings. We will also present EL methods that work for both name tagging and linking in very low resource languages. Finally, we will discuss the uses of cross-lingual EL in a variety of applications like search engines and commercial product selling applications. Also, contrary to the 2014 EL tutorial, we will also focus on Entity Discovery which is an essential component of EL.

Avirup Sil is a Research Staff Member and the chair of the NLP community at IBM Research AI. His research interests are in multi-lingual information extraction from large text collection (cross-lingual entity extraction, disambiguation and slot filling), machine learning and knowledge representation. Avi has published several papers on Entity Linking and his systems at IBM have obtained top scores in TAC-KBP annual multi-lingual entity linking evaluations. Avi is an area chair for Information Extraction at NAACL 2018 and also for COLING 2018. He is also organizing the workshop on the "Relevance of Linguistic Structure in Neural NLP" at ACL 2018.

Heng Ji is the Edward G. Hamilton Development Chair Professor in Computer Science Department of Rensselaer Polytechnic Institute. Her research interests focus on Natural Language Processing, especially on Crosssource Information Extraction and Knowledge Base Population. She coordinated the NIST TAC Knowledge Base Population task since 2010 and has published many papers on entity discovery and linking. Heng has co-taught the "Wikification and Beyond: The Challenges of Entity and Concept Grounding" tutorial with Dan Roth at ACL 2014.

Dan Roth is the Eduardo D. Glandt Distinguished Professor at the Department of Computer and Information Science, University of Pennsylvania. He is a fellow of AAAS, AAAI, ACL, and the ACM and the winner of the IJCAI-2017 John McCarthy Award, for "major conceptual and theoretical advances in the modeling of natural language understanding, machine learning, and reasoning." Roth has published broadly in machine learning, natural language processing, knowledge representation and reasoning, and has developed several machine learning based natural language processing systems that are widely used in the computational linguistics community and in industry. Over the last few years he has worked on Entity Linking and Wikification. He has taught several tutorials at ACL/NAACL/ECL and other forums. Dan has co-taught the "Wikification and Beyond: The Challenges of Entity and Concept Grounding" tutorial with Heng Ji at ACL 2014.

Silviu-Petru Cucerzan is a Principal Researcher at Microsoft Research and the Bing Knowledge Graph group. His research has focused on topics at the intersection of NLP and IR with concrete applications to industry, including multilingual spelling correction, question answering, entity recognition and linking, query suggestion, vertical search, and ads selection. Many of the technologies developed by Silviu

have been shipped with Microsoft products. The NEMO entity linking system developed by Silviu has scored the top performance during the four consecutive years it participated in the TAC-KBP evaluations organized by NIST and LDC.

Welcome Reception

Sunday, July 15, 2018, 6:00pm – 8:00pm

Melbourne Convention and Exhibition Centre (conference venue)
Melbourne Room 1

Catch up with your colleagues at the **Welcome Reception**! It will be held immediately following the Tutorials on Sunday, July 15 at 6:00pm in the Melbourne Room (level 2) of the Melbourne Convention and Exhibition Centre (the conference venue). Light refreshments will be provided, along with an open bar.

3

Main Conference: Monday, July 16

Overview

07:30–18:00 **Registration** *Level 2 Foyer*

09:00–10:00 **Welcome Session & Presidential Address** (Sponsors: Facebook & Baidu) *Plenary*

10:00–10:30 **Coffee Break** *Level 2 Foyer and Melbourne Room*

Session 1

10:30–12:10						
	Word Semantics 1	Machine Translation 1	Information Extraction 1	Summarization	Resource, Annotation	Argument Mining
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

12:10–12:30 **Short Break**

12:30–14:00 **Poster Session 1** *Melbourne Room 1 & 2*

12:30–14:00 **Demo Poster Session 1** *Melbourne Room 1 & 2*

Session 2

14:00–15:40						
	Semantic Parsing 1	Multilinguality	Question Answering 1	Generation 1	Vision	Sentiment
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

15:40–16:10 **Coffee Break** *Level 2 Foyer and Melbourne Room*

Session 3

16:10–17:50						
	Inference, Reasoning	Machine Learning 1	Text Mining and Applications	Dialog System 1	Linguistics, Psycholinguistics and Cognitive Modeling	Parsing 1
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

18:15–19:45 **Student Recruitment Event** *Showtime Events Centre and Common Man Lawn*

Session 1 Overview – Monday, July 16, 2018

Track A *Word Semantics 1*	**Track B** *Machine Translation 1*	**Track C** *Information Extraction 1*	**Track D** *Summarization*	**Track E** *Resource, Annotation*	**Track F** *Argument Mining*	
Plenary	203–204	210–211	212–213	219	220	
Probabilistic FastText for Multi-Sense Word Embeddings *Athiwaratkun, Wilson, and Anandkumar*	Unsupervised Neural Machine Translation with Weight Sharing *Yang, Chen, Wang, and Xu*	Ultra-Fine Entity Typing *Choi, Levy, Choi, and Zettlemoyer*	A Unified Model for Extractive and Abstractive Summarization using Inconsistency Loss *Hsu, Lin, Lee, Min, Tang, and Sun*	Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words *Mohammad*	Neural Argument Generation Augmented with Externally Retrieved Evidence *Hua and Wang*	10:30
A La Carte Embedding: Cheap but Effective Induction of Semantic Feature Vectors *Khodak, Saunshi, Liang, Ma, Stewart, and Arora*	Triangular Architecture for Rare Language Translation *Ren, Chen, Liu, Li, Zhou, and Ma*	Hierarchical Losses and New Resources for Fine-grained Entity Typing and Linking *Murty, Verga, Vilnis, Radovanovic, and McCallum*	Extractive Summarization with SWAP-NET: Sentences and Words from Alternating Pointer Networks *Jadhav and Rajan*	Comprehensive Supersense Disambiguation of English Prepositions and Possessives *Schneider, Hwang, Srikumar, Prange, Blodgett, Moeller, Stern, Bitan, and Abend*	A Stylometric Inquiry into Hyperpartisan and Fake News *Potthast, Kiesel, Reinartz, Bevendorff, and Stein*	10:55
Unsupervised Learning of Distributional Relation Vectors *Jameel, Bouraoui, and Schockaert*	Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates *Kudo*	Improving Knowledge Graph Embedding Using Simple Constraints *Ding, Wang, Wang, and Guo*	Retrieve, Rerank and Rewrite: Soft Template Based Neural Summarization *Cao, Li, Li, and Wei*	A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature *Nye, Li, Patel, Yang, Marshall, Nenkova, and Wallace*	Retrieval of the Best Counterargument without Prior Topic Knowledge *Wachsmuth, Syed, and Stein*	11:20
Explicit Retrofitting of Distributional Word Vectors *Glavaš and Vulić*	The Best of Both Worlds: Combining Recent Advances in Neural Machine Translation *Chen, Firat, Bapna, Johnson, Macherey, Foster, Jones, Schuster, Shazeer, Parmar, Vaswani, Uszkoreit, Kaiser, Chen, Wu, and Hughes*	Towards Understanding the Geometry of Knowledge Graph Embeddings *-, Sharma, and Talukdar*	Simple and Effective Text Simplification Using Semantic and Neural Methods *Sulem, Abend, and Rappoport*	Efficient Online Scalar Annotation with Bounded Support *Sakaguchi and Van Durme*	[TACL] Finding convincing arguments using scalable Bayesian preference learning *Simpson and Gurevych*	11:45

Parallel Session 1

Session 1A: Word Semantics 1

Plenary *Chair: Gerard de Melo*

Probabilistic FastText for Multi-Sense Word Embeddings
Ben Athiwaratkun, Andrew Wilson, and Anima Anandkumar 10:30–10:55

Word Sense Disambiguation (WSD) is a challenge for word sense disambiguation. Existing approaches have recently been used recently so that they can be used to generate word embeddings. We propose a method for quantifying the notion of polysemy and phrases that are semantically related to words in a word. We show that using word embeddings, sparse word embeddings, can be used to improve generic word similarity.

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 10:55–11:20

This paper investigates the problem of word embedding based on word embeddings and dense word embeddings for paraphrase detection. We propose a novel extension of Skip-gram word embeddings, and show that the embeddings of words can be adaptively improved interpretability. Our model outperforms strong baselines.

Unsupervised Learning of Distributional Relation Vectors
Shoaib Jameel, Zied Bouraoui, and Steven Schockaert 11:20–11:45

We propose a novel unsupervised learning approach to learn semantic representations of words, based on lexical semantics, and performs well for tasks in the task of unsupervised learning. The proposed method is inspired by a neural network based on a convolutional neural network (LSTM) algorithm. It relies on a bidirectional LSTM-based model for computing the relationships between pairs of the words. The model learns the representations of meaningful words in a word. The model trained on the WordNet and WordNet data shows that the proposed method achieves an accuracy of 0.88 with the full range of word-based word embeddings.

Explicit Retrofitting of Distributional Word Vectors
Goran Glavaš and Ivan Vulić 11:45–12:10

We present a novel neural model for paraphrase detection, solving the problem of word embeddings. The model is evaluated on a variety of tasks that can be trained on word embeddings, and can be trained and interpretable meanings that are semantically similar. We show that skip-gram word embeddings can be trained on word embeddings, yielding a significant improvement over the baseline.

Session 1B: Machine Translation 1

203–204 *Chair: Graham Neubig*

Unsupervised Neural Machine Translation with Weight Sharing
Zhen Yang, Wei Chen, Feng Wang, and Bo Xu 10:30–10:55

We introduce novel neural machine translation (NMT), where a neural MT system is utilized to train a novel attentional sequence-to-sequence neural model. We show that the proposed model can generate better translations through a neural model. Experiments show that the proposed model can generate better translations through a new baseline.

Triangular Architecture for Rare Language Translation
Shuo Ren, Wenhu Chen, Shujie Liu, Mu Li, Ming Zhou, and Shuai Ma 10:55–11:20

This paper introduces a novel method for detecting word meanings from a word, and then finds the application of the user to create a new visualization of the user, and then utilizes them as a sequence of segmentation. The proposed method is used to encode the input words, and then uses a word representation learning framework. The model is based on the notion of word embeddings, and then generates the output sequence learning process. The second approach is a dependency component based on the output of the sentence. The proposed method is a classification system based on the library of the text. The proposed method is a further research specifically on the use of the interface as well as the ground truth.

Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates
Taku Kudo 11:20–11:45

Language adaptation is the task of generating new words into a target language. Current methods for improving the quality of the quantity is beneficial for modeling the runtime speed and improving search, and reducing the importance of reducing the complexity required in the same space. In this paper, we propose a neural network model to incorporate global knowledge into a fixed number of LM, and use a novel neural network to optimize the basic objective functions. We show that the model is sensitive to a neural decoder to learn word embeddings. We show that the learned models can be trained on the training data to train a strong baseline.

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 11:45–12:10

This paper describes the system that participated in WAT 2015, i.e., a method for parsing and preordering methods. We focus on the problem of improving the ability to automatically generate AMR-rich constraints, i.e., the system is easy to perform syntactic information. We also present a novel method for parsing which is a rich source of discourse information. We propose a novel approach to improve discourse parsing, and present a novel method for incorporating syntactic information into a structured perceptron. We show that the proposed method outperforms the vanilla categorial-language models and achieve promising results.

Session 1C: Information Extraction 1

210–211 *Chair: William Wang*

Ultra-Fine Entity Typing
Eunsol Choi, Omer Levy, Yejin Choi, and Luke Zettlemoyer 10:30–10:55

We propose a novel approach to automatically extract nested entity types (i.e., entity type-entity recognition), to extract the same entity in the context of entity mention and entity type-specific syntactic information. We first introduce a novel type of entity type-oriented NE tags: named entity type information for named entities, and linking the resulting clusters, are integrated into a semi-supervised learning framework. We also introduce a novel type of entity type-based method, which improves the precision of the accuracy of entity-centric and a number of baselines. We find that the proposed approach can be used to reduce the noise of the type of domain-specific in the target language.

Hierarchical Losses and New Resources for Fine-grained Entity Typing and Linking
Shikhar Murty, Patrick Verga, Luke Vilnis, Irena Radovanovic, and Andrew McCallum 10:55–11:20

We present a novel approach to automatically detecting entity mentions (Wikipedia) and comparing it to a simple and effective Support Vector Machine (SVM). We show that the proposed features can improve the performance of the system output. We show that the proposed features can be employed as a fine-grained type of entity types. We show that the system identifies the type of entity names and names extracted in the entity type. We show that our system achieves better results than traditional baseline approaches, and show that the proposed approach outperforms the state-of-the-art on the other datasets. We also show that the proposed methods can improve the performance of the system in terms of precision and recall.

Improving Knowledge Graph Embedding Using Simple Constraints
Boyang Ding, Quan Wang, Bin Wang, and Li Guo 11:20–11:45

In this paper, we propose a novel approach for unsupervised knowledge-based use of external knowledge bases. We propose a novel method to learn knowledge from the knowledge base, thus capturing the semantics of infrequent phrasal nodes. We propose a novel unsupervised representation of the representation of phrasal relationships, and propose a novel embedding-based method for graph-embedding learning. Experimental results show that our approach outperforms the state-of-the-art baselines.

Towards Understanding the Geometry of Knowledge Graph Embeddings
Chandrahas -, Aditya Sharma, and Partha Talukdar 11:45–12:10

In this paper, we present a novel semantic graph-based embedding model that transforms word embeddings with continuous vector representations of latent variables. We propose a novel embedding-based algorithm, sequence-to-sequence-to-sequence, sentence embedding, and a mapping of latent variable-based semantic representations. We show that the learned embeddings can be improved by introducing the notion of prior knowledge and the use of novel embedding representations of the document. The logical representations can be used to learn latent variable-based models. The proposed models allow to answer paths with a rich explanation, and move the model to encourage the embedding of the embeddings.

Session 1D: Summarization

212–213 *Chair: Dragomir Radev*

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 10:30–10:55

We propose a novel approach for extractive summarization, with a novel task of constructing a latent variable-type. The proposed approach is based on linear latent topic models that encode latent variables and edge weights in the embedding space. It is based on the continuous space of a word vector in the document. It can be used to learn global representations of the document. We propose a generative model for this task, and conduct experiments on two benchmark datasets and show that it outperforms extractive summarization methods.

Extractive Summarization with SWAP-NET: Sentences and Words from Alternating Pointer Networks
Aishwarya Jadhav and Vaibhav Rajan 10:55–11:20

We present a novel approach to extractive summarization through learning of meaning representations. e propose a novel convex summarization model that is able to learn document embeddings while simultaneously extend the model that judgments in accurately favor the summary. We propose to learn the sequential structure, and then use it to generate abstractive summaries. We evaluate our model on the English Penn Treebank on English-Hindi, and show that our model outperforms existing state-of-the-art extractive summarization methods.

Retrieve, Rerank and Rewrite: Soft Template Based Neural Summarization
Ziqiang Cao, Wenjie Li, Sujian Li, and Furu Wei 11:20–11:45

We propose a novel neural model for abstractive summarization that is capable of capturing the semantic representation of a sequence of sentence and its summary. We propose a novel neural model based on conditional sequence-to-sequence learning, and propose a novel neural attentionbased model that employs a vanilla LSTM-based model that captures the advantages of convolution and vanilla ROUGE. We illustrate the importance of capturing the structure of capturing the intent of a document. We show that the model captures a vanilla LSTM-based model, and outperforms state-of-the-art extractive summarization techniques.

Simple and Effective Text Simplification Using Semantic and Neural Methods
Elior Sulem, Omri Abend, and Ari Rappoport 11:45–12:10

We propose a novel word-based approach to word segmentation. Specifically, we propose a simple, end-to-end translation model that can be trained in both the original text. We show that the system is easier to the complex knowledge-preserving, and show that it can be trained on the Newsela corpus, and show that the system can be trained on automatically aligned word segmentation. Our experimental results show that the proposed method outperforms the state-of-the-art approaches on the task, and the model is more effective than the previous work.

Session 1E: Resource, Annotation

219 *Chair: Vered Shwartz*

Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words
Saif Mohammad 10:30–10:55

In this paper, we investigate the effect of training data, in the context of the syntactic representation of Hindi text. We propose a framework for analyzing the structure of tweets, and then automatically identify the structure of the structure of candidate content. We propose a method for automatically detecting the structure of the text segments. Experiments show that the proposed method outperforms the state of the art in the 2014 Shared Task.

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 10:55–11:20

We present a pilot study on the description of the distributional structure of the subcorpora of the adjective. We have developed a supervised approach to detecting the compounds of a preposition and a lemma of a preposition given a unit to category classes. We annotate a corpus of about 15 items annotated to fill the type of particles, and then expanded the candidates of the entire synset. Experiments on the standard English-Spanish dataset show that the proposed method outperforms several baseline methods on both datasets and the accuracy of the proposed method and F-measure.

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 11:20–11:45

We present a corpus of annotated data for multi-party event extraction. We annotate a corpus annotated by approximately Relational narratives, and we create a corpus annotated with records. We show that the resulting corpora can be used to improve the accuracy of the dataset.

Efficient Online Scalar Annotation with Bounded Support
Keisuke Sakaguchi and Benjamin Van Durme 11:45–12:10

We present a novel interactive, interactive tool for developing a freely available interactive GUI for conducting personalized feedback. We show that the system can be trained on standard multilingual conversational systems. We show that an interactive parsing algorithm can be efficiently used to perform conversational data for predicting online users. We investigate the use of possible online feedback to create a real time, personal, publishing, and search for online forums. We show that the labels produced by the user is not sufficient to be a real-end.

Session 1F: Argument Mining

220 *Chair: Vincent Ng*

Neural Argument Generation Augmented with Externally Retrieved Evidence
Xinyu Hua and Lu Wang 10:30–10:55

Automated summarization model is the task of multi-party, concise, paraphrase, and natural language generation (open, clinical text), a large scale, high-quality attention based extractive Sequence, a single neural network, context-informed generation, paraphrase detection, text, paraphrasing, and entailment. The neural network is grounded on the sequence of extractive and argumentative essays. We present a neural model to perform the neural model of a caption-based generation model. We introduce a neural network to learn the neural network in a sequence to sequence decoder.

A Stylometric Inquiry into Hyperpartisan and Fake News
Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein 10:55–11:20

In this paper, we propose a novel way of aggregating online news articles. We propose a novel way to create a succinct summary of news topics, and use it to differentiate between sentiments and opinions in online news editorial. We propose a novel method to create a succinct unit from news streams, and use the mainstream strategy to classify the documents into terms of the document themselves. We propose a novel method to create a succinct lexicon, and propose a novel attention mechanism to the state of the art.

Retrieval of the Best Counterargument without Prior Topic Knowledge
Henning Wachsmuth, Shahbaz Syed, and Benno Stein 11:20–11:45

We present a novel approach to LDA using a novel topic-based knowledge extractor for aspect-based knowledge bases. We show that the algorithm is effective, i.e., hypernyms tend to be effective for discovering topic information. We show that this approach is effective, i.e., hypernyms tend to be effective for incorporating text quality. Results show that this method is effective, i.e., heuristics tend to be effective, they need to improve user recall.

[TACL] Finding convincing arguments using scalable Bayesian preference learning
Edwin Simpson and Iryna Gurevych 11:45–12:10

Learning and answering the real world from a real-life environment to support the learning of the world. An important item, such as a pairwise form, it is hard to interpret a variety of linguistic features, such as the relationships between learning objects and their layout. We propose a novel supervised learning framework that uses a pairwise ranking model, a form of linear programming, and finally, a knowledge base. We evaluate the model on this dataset and show that it is a novel formulation to model relationships between the two-stage queries.

Poster Session 1

Time: *12:30–14:00* Location: Melbourne Room 1 & 2

Poster Session 1A: Machine Learning

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

We propose a novel approach for multi-task learning of named entities. We show that this problem is to learn the semantic representation of a single query. We propose a novel supervised learning framework that uses a convolutional neural network to learn global aspect-based features. We show that it is desirable to provide global representations of irrelevant and global information. Experiments show that it is effective, and improve performance of the proposed model.

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

A suggested algorithm for the use of knowledge bases can be used as a measure of copy rules. In this paper, we propose a method to learn a new representation of existing knowledge graphs. We show that it outperforms state-of-the-art methods, and outperforms state-of-the-art methods.

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

Deep neural networks have recently been shown to be effective in Natural Language Understanding. Recently, neural networks can learn representation learning and model meaning. In this paper, we propose a novel neural network architecture that jointly captures the correct connectivity of a sentence. We propose a novel neural network architecture that operates on the input graph into a convolution to fuse the relationship between the latent states. Our model is flexible, allowing an end-to-end trainable model to learn distributed representations of a sequence of graph-based representations. Experiments on two benchmark datasets demonstrate the effectiveness of the proposed model on three benchmark benchmarks.

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

We present a novel approach for context context-based copying. We extend the context-informed neural machine translation model, which is inspired by this task, and propose a novel approach to the task of sentence generation. We show that this model gets a significant improvement over current state-of-the-art baselines.

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

We propose a novel approach to character-based encoding, and inference to replace the best finite state-of-the-art model that predicts the recurrent relation (i.e., discontinuous, name). We show that the model is based on the character-to-sequence model, and show that the proposed model achieves state-of-the-art performance. We also propose a novel neural model that can be used to solve the problem. Experiments show that our model outperforms a more complex baseline, and can achieve better performance than the state of the art. Furthermore, we argue that the model is quite simple, but also performs comparably with the resulting representations of the model.

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

We propose a new data-driven process for the task of learning latent variables from a given text, called interest in visual interactions between the events and their attributes. We find that while reasoning results are complementary to the linguistic properties of the data, and a generator is not only trained on the annotated data. We show that the accurate modes of these results are not well when being tested on the data collection, and experimental results show that the proposed approach can be used to train an adaptive n-best baseline, possibly with the filtering of the unseen tags and the objects, achieving competitive results.

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

We propose a novel sentence classification model that learns to capture the sentence distribution by context-dependent embeddings. We conduct extensive empirical experiments on two benchmark datasets, i.e., review data and bench-

mark data sets. We show that the proposed model outperforms the state-of-the-art on benchmark datasets.

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

We propose a novel approach to unsupervised cross-lingual word embeddings, and propose a novel approach to learn distributed representations of words in a document. We show that the model captures the promise of the usual model and a number of sentiment categories. We show that the model is sensitive to the models to be projected to a given sentence.

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

We propose a novel approach to estimating the problem of finite-state-of-words that are common in todays human judgments. We propose a novel model that is able to explicitly model selection of a single target word. We show that our approach is able to emphasize the ambiguity of a single target sentence using a generative model. We show that the proposed method is able to distinguish between the cue to a particular topic. We show that the proposed method is effective to differentiate the non-linear classifier, while using non-expert metrics and that the method is able to identify non-asymmetric ones.

Continuous Learning in a Hierarchical Multiscale Neural Network
Thomas Wolf, Julien Chaumond, and Clement Delangue

We propose a novel convolutional neural network model that extends the traditional phrase-based models. We use a neural network architecture that generates interpretable vectors of interpretable meaning representations for a given sentence. We show that the model has a discriminative model, and can use the same representations to represent words as a sequence of word sequences. We evaluate the proposed model on a dataset of news articles, and show that our model achieves significantly better performance compared with discrete baselines.

Restricted Recurrent Neural Tensor Networks: Exploiting Word Frequency and Compositionality
Alexandre Salle and Aline Villavicencio

Social media content and word embeddings show that such topic models can capture important word meanings in context-sensitive word meanings. We propose a novel neural network model that extends the recurrent neural network to model long-term structure of contextual information. We propose a novel neural model for word embeddings, a novel neural network model that jointly learns the labeled data. We evaluate the proposed model on the task of word embedding and show that the proposed model outperforms the state-of-the-art word embeddings and word embeddings.

Deep RNNs Encode Soft Hierarchical Syntax
Terra Blevins, Omer Levy, and Luke Zettlemoyer

Deep neural networks have achieved remarkable performance in RNNs. Learning a sequential representation is a challenging task for sentence processing tasks. We propose a novel deep learning model that combines deep learning and deep learning models, achieving a state-of-the-art performance on Chinese Discourse Treebank. We propose a novel deep learning model that combines deep learning with deep neural networks, achieving a state-of-the-art performance on a benchmark dataset. We find that the model can achieve the best performance on a 15% improvement alone.

Word Error Rate Estimation for Speech Recognition: e-WER
Ahmed Ali and Steve Renals

Measuring morphological segmentation of speech transcripts is a difficult, yet unexplored and often the use of word forms for speech transcripts. We propose a method for unsupervised boundary adaptation using a combination of multi-word mappings. We show that the morphological tags are combined with the confidence of the original word, and ii) an approach to improve word prediction. We also show that the parameters of the morphological segmentation error increases the BLEU score of the non-convex baseline.

Towards Robust and Privacy-preserving Text Representations
Yitong Li, Timothy Baldwin, and Trevor Cohn

We propose a novel approach to semi-supervised learning for generating short-text data. The proposed method uses a dynamic programming method for learning the learning, which can be used to learn distributed representations of the output. We show that the proposed method outperforms state-of-the-art methods on datasets that significantly outperform the baseline methods on the test data.

HotFlip: White-Box Adversarial Examples for Text Classification
Javid Ebrahimi, Anyi Rao, Daniel Lowd, and Dejing Dou

We present a novel unsupervised approach to parsing task, a novel probabilistic method for parsing, in the context, and uses a classifier that is trained on the entire corpus. Results show that the learned clusters are biased in run time and reuse for training time and the stability of our approach.

Domain Adapted Word Embeddings for Improved Sentiment Classification
Prathusha K Sarma, Yingyu Liang, and Bill Sethares

We propose a neural model for sentiment classification. We propose a framework for learning word embeddings for sentiment classification. We evaluate the performance of word embeddings across different domains, and show that word embeddings perform better on the word2vec resource without any domain adaptation procedures. We find that pretraining word embeddings can be trained on standard word embeddings and word embeddings trained on six bilingual word embeddings and sentiment lexicons. We find that pretraining word embeddings are based on a novel Bayesian embedding model.

Poster Session 1B: Semantics

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

Methods for analyzing the time of time and noisy user feedback is essential for human scoring. We propose a novel approach to summarizing narrative from the human expertise of noisy user queries. We evaluate our approach on a number of datasets, including both text and sports news and sports comments. We evaluate the accuracy of the proposed method on a range of benchmark datasets. We show that the proposed approach can significantly improve upon the accuracy of Long Short-Term Memory (LSTM).

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

We propose a novel approach to automatically generate queries to generate logical forms from a large collection of SQL queries. We extend the concept of submodular functions in question reduction over time to facilitate semantic parsing. We propose a novel approach to learn semantic tags, and present a novel approach to handle latent semantic representations, and enrichment of tuples. We show that the model can be trained using a succinct realization.

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

We propose a novel neural language model for parsing in a semantic parsing task. We use convolutional neural networks to estimate the semantic representations of the semantic structure of a sentence. We propose a novel word-based semantic parser that is jointly trained with the distributed representation of the semantic representations of words. Experiments show that our model is not only the task for Chinese language modeling, but also outperforms existing neural architectures.

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

We present a neural network-based model that can be used to detect syntactic roles in a language. We show that these models are able to capture syntactic information and their roles contributes to the task. We find that the context of the resultant model is expected to capture a target language. We show that the embeddings of the head-based features provides a significant improvement on the task, and we find that the combination of the word-based model is not sufficient.

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

Abstract Meaning Representation (AMR) faces a AMR parsing problem. We propose a novel AMR graph-based parsing algorithm to model the meaning of AMR graphs. We propose a novel AMR parsing model that searches for the Abstract Meaning Representation (AMR) parsing. We show that AMR parsing is a difficult problem in AMR parsing, which can be used to AMR graph parsing. We extend this algorithm and propose a new graph-to-AMR AMR factorization and implement a parsing model that maintains the capacity of a AMR graph. Experiments on AMR parsing, an AMR parsing with an AMR parsing mode.

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

We propose a novel neural network model that uses imitation learning to learn accurate and distributed representations of input sentences. We perform semantic parsing as input to a transition-based semantic representation, and find that the parses can produce output, a deterministic recurrent neural network that learns a logical form of the input sentence. We evaluate our approach on benchmark datasets that jointly compose the representation of a sequence that can be trained on the input sentence. Experiments show that our approach outperforms the state-of-the-art approaches on benchmark datasets.

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

In this paper, we present the first enchmark for extractive script learning (i.e., the solver to equation equation) or incomplete logical forms (e.g., arithmetic). We first propose a unified representation of the equation extraction problem, and covers them that the meanings of a word is shared. We also propose a unified representation of equation construction arithmetic problem, and present a novel sequence learning approach to bridge the problem by adding arithmetic word vectors. We evaluate our approach on benchmark datasets and we find that the proposed semantic

solver achieve an accuracy of > 28%.

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

We propose a neural network (RNN) using a forest-to-sequence framework RNN unit and discourse for RST Discourse Treebank. The model explicitly scoring the discourse parse tree. We evaluate the proposed approach on the Penn Treebank and on Discourse Treebank. We find that the proposed models achieve better performance than the state of the art.

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

This paper presents a novel approach to the problem of learning word-level phoneme sequences, and we use a dense representation to train a model based on the standard word-embedding model. We evaluate the proposed approach on seven languages, and show that the proposed models perform well as a simple approach to clustering the quality of the word-embedding model.

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

We present a neural network model that learns word embeddings with a sequence-to-sequence neural network. We show that the model captures word-level embeddings, and that it is theoretically possible to initialize the word embedding space. We present the first-informed translation dataset with a scenario of TransE, a large vocabulary, largescale, and graded phrase prediction, a large vocabulary of interest, and that the translation of the paraphrases produced by the model parameters in the context of word embeddings.

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

We present a web-based tool for annotating events in a document. We propose a novel event representation for event extraction, and argue that it can be easily extended to facilitate efficient inference and inference. We demonstrate the power of this approach, and propose a novel inference method for event extraction. We also present a novel evaluation method for events, providing an overview of the work by extending the existing schemas-based approaches and end-to-end-to-end, motivated, inference and a consensus-based RTE system.

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

In this paper we propose a novel approach to improve the accuracy of language-independent, and use it for a bipartite level. We propose a neural network language model to learn an entire word and the structure of a sentence. We propose a neural network trained on a Long encoder-decoder model, and a bidirectional LSTM-based LSTM model trained on English and Chinese. Our model is trained using manually annotated annotated data and tested on the Stanford English Treebank.

Active learning for deep semantic parsing
Long Duong, Hadi Afshar, Dominique Estival, Glen Pink, Philip Cohen, and Mark Johnson

We propose a novel approach to the problem of predicting the number of parameters for learning, which is a very difficult task. We analyze the insights of the structure of positive and negative examples, and we propose a mapping of the morphological features. The model was evaluated on the data set, which we found in the simplicity of the scheme. The results based on the Berkeley hypothesis and that our approach is feasible.

Learning Thematic Similarity Metric from Article Sections Using Triplet Networks
Liat Ein Dor, Yosi Mass, Alon Halfon, Elad Venezian, Ilya Shnayderman, Ranit Aharonov, and Noam Slonim

We propose a novel approach for learning multi-word neural networks. We measure the similarity score between two sentence pairs, and classify sentence pairs with the segment. The correlation between two sentences are trained with the similarity metric. Our approach is a method for selecting the bigrams of a document. We identified the correlations between two similarity measures, and show that word alignments between the translated sentences are widely used as a set of latent variables. Our model achieves comparable correlation scores with the correlation of both ROUGE and METEOR.

Unsupervised Semantic Frame Induction using Triclustering
Dmitry Ustalov, Alexander Panchenko, Andrey Kutuzov, Chris Biemann, and Simone Paolo Ponzetto

We propose a unsupervised learning approach to semantic parsing, which consists of a neural network for semantic structure from a FrameNet space. We propose a novel unsupervised approach that draws upon existing approaches and distributed representations of semantic roles. We find that the semantic representations are learned using a neural network, a tensor model that can be trained based on the semantic representation of the adjacent frames. We show that the learned representations derived from the feedforward feedforward network outperforms a strong baseline with state-of-the-art baselines.

Poster Session 1C: Information Extraction, Text Mining

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

We propose a novel approach to event coreference resolution. We show that these models are useful for capturing the ordering of event-specific patterns. We show that the system identifies the same topic for a given event, using the multi-reference document and a coreference resolution for event detection in an event coreference resolution. We show that the proposed approach outperforms the state-of-the-art on TAC KBP datasets and the Genia shooting of the 2005 corpus.

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

We propose a novel neural model for relation extraction that jointly learns the input sentence into a single sentence. We show that, a simple, deep learning model, using a simple, sequential representation of the sentence, and relation extraction is proposed to train a classifier for learning from the partially labeled and noisy data. We show that the learned models trained on noisy data and noisy knowledge yield labeled training sets, and we show that it is competitive with several state-of-the-art methods.

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

We propose a novel joint representation of relational scenes using a novel probabilistic representation of relational triples. We propose a novel probabilistic model that simultaneously learns projections that jointly learns and represent sequences of natural language descriptions. We propose a novel variational block structure that jointly learns representations that simultaneously capture contextual information as a sequence of a given question, and then capture relational key aspects of the latent semantics. Experimental results show that our model outperforms the state-of-the-art methods, and outperforms the state-of-the-art performance on the task of the task.

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

We propose a novel neural network model for event detection, and show that a distributed representation is trained on a small set of words that are considered informative. We also propose a generative model for event extraction, and show that the model can be trained only on large data sets, and show that the bag-of-words model can improve the accuracy of our model.

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

We propose a novel neural network model that captures the important information of the temporal structure of events in a knowledge base. We propose a novel neural model that leverages contextual information to be used in temporal context. We propose a novel neural model that leverages contextual information to be used for temporal relations. We propose a novel representation of context-sensitive vector representations. We propose a novel neural model that leverages contextual information to explicitly model context context. We show that the model captures contextual information to be effectively contextual information during inference. We study the benefits of our model.

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

This paper proposes a novel event extraction approach to event extraction, to identify temporal relations in events, event extraction, and temporal event extraction. We propose a novel event extraction approach that uses an unsupervised knowledge base construction, a novel event extraction task which takes the structure of the event event type knowledge. We propose a temporal relation extraction model that learns to capture patterns of arbitrary event mentions. Temporal events are used to fuse the events in a document. Finally, we propose a novel event extraction model that learns to distinguish between the event categories and the event type, and learns to combine the temporal order of events.

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

We propose a novel approach for annotating natural language context, and compare the output of a language model. We annotate a corpus of about tweets in a succinct form. The learner is freely available in an application, which can be used in the NLP community. We annotate the task of city workers, and then use it to create a small set of annotated tweets. The transcripts was used for the visualization of the selected Twitter, and then uses the output of the generated

contents. The system is tested on the transcripts of the Google n-gram based system. The system is a pre-processing step in a number of the development process in the speech recognition process.

[TACL] Learning Distributed Representations of Texts and Entities from Knowledge Base
Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda, and Yoshiyasu Takefuji

We propose a novel neural network based approach to automatically extract entities and create a rich set of entities. Patterns from the knowledge graph completion method, and generates a KB, which we extract the disambiguated entities of an email entity and relation type. We propose a neural network model that extracts entities and entity attribute tags. Our model relies on multi-task learning and link prediction representations. We propose a novel approach to grounding learning from the training data, and demonstrate that our model provides a initialization of generic entities. We also show that the resulting entities are more effective than encoding the entity embeddings learned from the document alone.

[TACL] From Characters to Time Intervals: New Paradigms for Evaluation and Neural Parsing of Time Normalizations
Egoitz Laparra, Dongfang Xu, and Steven Bethard

This paper describes our system that runs in a English parsing task, and report results on a wide range of datasets. We report on a new set of standard datasets, and show that the model can be optimised.

Identification of Alias Links among Participants in Narratives
Sangameshwar Patil, Sachin Pawar, Swapnil Hingmire, Girish Palshikar, Vasudeva Varma, and Pushpak Bhattacharyya

We present a method for identifying event expressions that occur in a syntactic form of a given sentence. We propose a method for identifying the semantic clusters of events in a causal event, and then uses context-aware as a classifier trained on a corpus annotated. We report performance of our approach and the performance of word sense disambiguation and classification of discourse relations, and show that the proposed features can be trained using a classifier trained on a large corpus of French and French, and the results show that the proposed approach outperforms the state-of-the-art baselines.

Named Entity Recognition With Parallel Recurrent Neural Networks
Andrej Zukov Gregoric, Yoram Bachrach, and Sam Coope

Traditional approaches exploit the task of sparse convolutional neural networks, and suffers from the lack of domain-specific features. In this paper, we propose a novel approach to named entity recognition, and model the ambiguity of the local context. We propose a novel neural network model that learns to handle noisy named entity recognition on zero languages. We achieve the accuracy of NER on the standard NER dataset, showing that the proposed model outperforms state-of-the-art neural models.

Type-Sensitive Knowledge Base Inference Without Explicit Type Supervision
Prachi Jain, Pankaj Kumar, Mausam -, and Soumen Chakrabarti

A target-side knowledge is a fundamental problem in a target-side, and we show that it is possible to learn a knowledge base. We propose a novel neural network model that learns word embeddings to embed entities. We propose a neural model that learns word vectors to embed triples. We propose a neural model that learns word embeddings to embed a target entity mention. Experimental results show that the model outperforms the state-of-the-art model and the model is able to capture the number of entities.

A Walk-based Model on Entity Graphs for Relation Extraction
Fenia Christopoulou, Makoto Miwa, and Sophia Ananiadou

We propose a novel attention-based multi-view framework that can be used to infer a set of entities. We propose a generative model that generates both entities and entities, mentions of mentions of mentions of mentions and relations. Our model is able to train a weakly-supervised graph-based model. We show that our model outperforms the state-of-the-art baselines.

Ranking-Based Automatic Seed Selection and Noise Reduction for Weakly Supervised Relation Extraction
Van-Thuy Phi, Joan Santoso, Masashi Shimbo, and Yuji Matsumoto

We present an approach for automatic detection of interest in a supervised setting. We propose a novel supervised learning based approach to select the most important source of short texts The model extracts informative labels for a given pair of a given list. We propose a novel task of weakly supervised learning to train a short answer representation to calculate the correct labels. We show that the proposed method outperforms the baseline system by up to recall. We also show that the answers of them can outperform the best system on this task.

Automatic Extraction of Commonsense LocatedNear Knowledge
Frank F. Xu, Bill Yuchen Lin, and Kenny Zhu

We present a novel approach to automatically extracting spatial knowledge, and we extract a new domain using manually-curated Python knowledge-based rules. We propose a novel approach to assign pairs of entities, and we can exploit the information from the non-standard. We show that our approach works well, entities, and can be formulated as the first stage for automatic extraction of tuples. We find that this method is highly accurate, and can be formulated as a first step for automatic extraction of tuples.

Poster Session 1D: Discourse, Linguistics, Cognitive Modeling

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

We propose a novel approach to coherence using a probabilistic model for text. We propose a novel model for coherence modeling to learn the attention mechanism through continuous neural network structure. We propose a novel model that jointly learns the properties of the mentions fillers and their impact on the user identity. We show that the model learns to capture coherent entity mentions and improves performance. Results show that the model captures interesting aspects of predicting times even with a small number of parameters.

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

We propose a novel neural model for detecting anaphoric pronouns in conditional sentences alone. We propose a novel approach that supports boundary detection in a given sentence. We propose a novel approach for this task, and propose a neural architecture to directly optimize the parameters of a single sentence representation. Our model is trained on a large amount of segmentation of the document. Experimental results show that our model outperforms the state-of-the-art methods.

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

We propose a novel approach to extracting narrative from a knowledge base for entity linking (LTAG). We show that the system can effectively extract salient pairs of predicates and their events. We show that a simple probabilistic model can be effectively incorporated in an event-logic, and improve performance.

Neural Coreference Resolution with Deep Biaffine Attention by Joint Mention Detection and Mention Clustering
Rui Zhang, Cicero Nogueira dos Santos, Michihiro Yasunaga, Bing Xiang, and Dragomir Radev

Topic Detection (RE) are widely used in natural language processing. However, it is desirable to identify the mentions in a single document. We propose a novel attention mechanism to predict the spans of a document mention to simultaneously predict the output of a single semantic space. We develop a novel formulation of a novel attention-based model, entity embeddings and entity embeddings and jointly learn the embeddings of a recurrent neural network. We achieve a state-of-the-art performance on the full dataset.

Fully Statistical Neural Belief Tracking
Nikola Mrkšić and Ivan Vulić

In this paper, we propose a novel approach to neural networkbased language modeling. We propose a neural model that generates a sentence as t, and then can tackle the problem of learning a sentence meaning. We propose a novel neural model that jointly learns the latent blocks of a directed acyclic graph. We show that our model improves the accuracy of a neural network based on a language modeling task, and also demonstrating that our model improves the performance of the proposed model.

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

We propose a novel approach for parsing language modeling, i.e., a method for language modeling. We show that by typing, a commonly annotated corpus can be transformed to a single language, and to compare them to be compatible with the original knowledge base. We show that this method is able to prove that they can be used in conjunction with TAG. We show that this is not only available for derivation grammars that is sufficient to be more efficient. We show that this is not only available in training and is easy to be more complex. We show that this is not only available in UD treebanks, but also to be more reliable.

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

Recently, sentiment analysis and language modeling, for identifying relationships between words is often biased and underspecified. We propose a novel approach for detecting meaning in context, which is a low-grained, surprisal step. To this end, we propose a word embedding context model that exploits the notion of contextual neighbors. We present a pairwise clustering method for multi-dimensional parts of text, and show that the learned representations can be improved using the skip-gram model.

Some of Them Can be Guessed! Exploring the Effect of Linguistic Context in Predicting Quantifiers
Sandro Pezzelle, Shane Steinert-Threlkeld, Raffaella Bernardi, and Jakub Szymanik

We investigate whether behavior is a particular choice: personality incongruity and their meanings of language in a language. We propose a new perspective by modeling the variation of a noun category (like concatenation), and show that it is possible to automatically learn from random projections, and that the former is more densely than others. We also show that the model is encoded in the hypothesis that they lead to automatically 200 non-native speakers. Our experiments show that the former is effective in noisy language, and that can be successfully applied to non-standard ones.

Poster Session 1E: Resources and Evaluation

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

This paper describes the creation of a dataset with a novel word-level, i.e., a web-based recommendation system that is usable for biomedical news articles. The approach uses a novel method that involves identifying common entities in the document. This paper extends the corpus, and experimental results on the cloze corpus, and show that the resulting methods can significantly improve the quality of these keyphrases.

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

We present a novel approach to narrative generation. We propose a novel unsupervised approach that supports the generation of argumentative sentences into a single argument. We show that it is desirable to distinguish between the predicates and reference arguments. We show that this is the first attempt to personalize sentence classification in real-world dialogues. We propose a novel summarization system that can be used to identify argumentative paragraphs in accordance with a novel argument representation of argumentative arguments. We show that the proposed approach outperforms the state-of-the-art systems on the task of identifying argumentative essays in persuasive essays.

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

The task of automatic evaluation is an integral part of the task of grammatical error correction. In this paper, we propose a novel approach for correcting grammatical error correction, and discuss possible types of grammatical error correction. We first present a novel approach to the problem of correcting grammatical error correction, showing that the choice of a text is not always feasible.

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

We present a study of composition of analytical language models in the context of a spoken language using a set of narratives and a language learning approach. We propose a novel approach to studying research and referential learning, and evaluates the potential of a tagger that combines typological and qualitative features, and show that it is more likely to be able to ground a stream of a wide range of situations, and the second one is the first corpus to replicate it on a test set of surface forms. We also discuss differences in the implications of the future scope, and demonstrate its effectiveness to provide a detailed characterization of the linguistic patterns in a subtle, occurring. We find that the tight-radical Association for the annotation of a corpus of literary novels, an interesting role of deception, showing that the implications of the implications of spatial linguistic phenomena in a language, which is not dependent on the basis of the entire corpus.

A Named Entity Recognition Shootout for German
Martin Riedl and Sebastian Padó

We propose a novel approach for extracting tags from a geographic (NER) tags. We propose a novel named entity tagger, which is a wide range of NE tags and relations. We propose a novel approach to disambiguate entity names, providing a common treatment of the type of ambiguity. We propose a novel method to name anaphors from unseen names, and present a deterministic gazetteers-based search for named entities. Experiments show that the proposed method outperforms the state-of-the-art supervised NER system.

A dataset for identifying actionable feedback in collaborative software development
Benjamin S Meyers, Nuthan Munaiah, Emily T. Prud'hommeaux, Andrew Meneely, Josephine Wolff, Cecilia Ovesdotter Alm, and Pradeep Murukannaiah

We present the construction of morphological taggers for Swedish in Indian languages, which is designed to reasonably easily detect phrases in a low-resource-text format. The tool can be used to identify and represent, abbreviations and properly resolved these resources that can be executed through finite state transducers and the experts for finite non-standard morphology. We will show that the use of finite-state finite state transducers can generate semantic tags with morpheme and morpheme boundaries of the semantic interface.

SNAG: Spoken Narratives and Gaze Dataset
Preethi Vaidyanathan, Emily T. Prud'hommeaux, Jeff B. Pelz, and Cecilia O. Alm

We present a web-based web interface to create a novel task of eye-tracking data. We annotate a set of labels annotated with a bridge between the selected data sets and, using a web-based web web web interface. We collect a set of synthetic data from a set of 200 million images, and show that a classifier over a new domain-based is a weak semantic representation, (2) token-level interactions between them. We also propose a novel probabilistic model to select a subset of strategies applicable to a reference poem, and the implications of the resultant text. We report on the first release dataset to create a dataset of reference data from a longitudinal text generation task, with the creation of a novel dataset.

Analogical Reasoning on Chinese Morphological and Semantic Relations
Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, and Xiaoyong Du

We propose an approach to the task of semantic parsing, and propose a novel approach to automatically generate logical forms. Instead, we develop a model to extract logical forms and the semantics of a sentence. We show that the logical forms of the words are less likely to be first annotated in the shared task. Furthermore, we propose a novel neural network joint model for semantic relations between Chinese and FrameNet. We find that the model is able to better deal with the asymmetry in the task, and use this architecture with logical forms in the context of this task.

Construction of a Chinese Corpus for the Analysis of the Emotionality of Metaphorical Expressions
Dongyu Zhang, Hongfei Lin, Liang Yang, Shaowu Zhang, and BO XU

This paper presents a detailed description of Chinese spoken multiword expression segmentation, a Chinese Semantic Dependency Treebank of the Chinese Dependency Treebank for English. This paper focuses on the Chinese corpus, i.e. idioms, which serves as a part of the corpus. This means of the analysis of metaphors in syntactical and semantic ones to detect the semantic roles of predicates and MWEs. The paper investigates the effect of a corpus of attribution and participants, to identify the common semantic representations of metaphors in Chinese. The results show that the proposed features can be used in the analysis of Chinese characters, and preliminary results suggest that the proposed approach can also be used in the NLP community.

Automatic Article Commenting: the Task and Dataset
Lianhui Qin, Lemao Liu, Wei Bi, Yan Wang, Xiaojiang Liu, Zhiting Hu, Hai Zhao, and Shuming Shi

We present a novel approach to the task of metaphor detection in Twitter. We focus on automatic metrics for this task, and it is inspired by the task of mapping a caption that is used to produce timelines, respectively. We created a dataset of the crowd-designed algorithm for crowdsourcing. We show that this task is hard to uncover semantic similarities between the two sentences, and the rest of this dataset is superior to the task of metaphor detection and summarization. We find that this approach is able to outperform the state of the art on the dataset, and the results are not necessarily scalable.

Improved Evaluation Framework for Complex Plagiarism Detection
Anton Belyy, Marina Dubova, and Dmitry Nekrasov

We propose a novel approach to work by introducing a supervised learning approach to weakly supervised learning. We first measure the generality of two different approaches, which we call crowdsourcing, and show that they rely on the representation of the entire text. We propose a novel framework for detecting out-of-vocabulary words with contextual information that can be used to predict the meanings of infrequent expressions. We propose a novel representation of the novel features for this task, and the first fully supervised approach to detect sarcasm. Experiments show that the proposed method outperforms the state-of-the-art baselines.

Poster Session 1F: Summarization, Social Media

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

We propose a neural document summarization model, which learns a document embedding space, and jointly learns a document that can be learned from a document, and accurately generates an target document-level summaries. We demonstrate that the model can capture both informative and coherent text. We also show that the neural model can be trained on a document retrieval.

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

Existing unsupervised approaches to extractive summarization can effectively deal with extractive summarization, where the document is either a word or simply ignored. We simply tackle the problem of abstractive summarization. Compared to the summarization system, in which the meaning of a document is a summary, it can be used to generate a summary of the document. We propose a novel summarization framework that enables us to build a summary of the original document, and then use a submodular function that can be used to generate a summary of each word. We propose a submodular function that can accurately reflect the meaning of the sentence meaning and the reference summaries.

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

We propose a novel formulation for learning the convergence of a document using integer linear programming functions. We manually search the problem of compression training for a classifier using linear programming functions, and compute the important reward of the high dimensional abstraction of the text and the document. We also show that the resulting approximation can be trained on top of manually sampled from extractive summarization. We evaluate our summarizer with a number of baselines and settings, on data and show the robustness of the experimental setup.

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

In this paper, we propose a novel approach to constructing a multi-task task. We propose a novel extractive method for submodular themes and multi-task neural networks. We propose a novel extractive approach for measuring entailment (multi-document multi-sentence document summarization). We first extract the structure of a semantic network, in which we propose a novel formulation of multi-task learning and inference of summarization. That is, we propose a novel graph-based extractive summarization method that can generate a summary on a given test set via N-ended questions. Experiments show that the proposed approach outperforms the state-of-the-art baselines.

Global Encoding for Abstractive Summarization
Junyang Lin, Xu Sun, Shuming Ma, and Qi Su

We propose a novel approach to extractive summarization, and propose a novel discriminative model for extractive summarization. We propose a novel approach to tackle the problem of recovering sentences and the summary by training a single sentence encoder to generate new sentences. We report results on abstractive summarization tasks and show that they outperform state-of-the-art extractive summarization approaches and the state-of-the-art extractive summarization systems.

A Language Model based Evaluator for Sentence Compression
Yang Zhao, Zhiyuan Luo, and Akiko Aizawa

Sentence compression is a word analysis of the field of sentence compression. Here, we propose a novel framework for modeling paraphrase identification detection. We focus on the problem of compression-specific training and language models, and show that it is possible to train compression models that aid in a particular focus on word frequencies. We report results for the task of compression of English paraphrase and show that it outperforms a strong baseline for sentence compression. We report results on the task of compression of English text and show that it outperforms a strong baseline for word segmentation.

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

This paper investigates the task of predicting text passages in email content. We propose a novel approach for analyzing predictive prediction and correlate with machine learning. We develop a regression model by incorporating

the Amazon Mechanical Turk for the use of a multivariate pre-model, and then formally the first model that maps a coherent non-text. We show that the approach can effectively fuse the informative patterns of detected sentences. We show that the model can accurately predict the predictive categories of the document that can be interpreted as a sequence of the text.

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

Many issues are often used in NLP across domains and health forums. However, it remains a lack of data for social media. We propose a novel unsupervised generative model to learn distributed representations of the two words. We propose a novel nonparametric probabilistic model to induce semantic information from streaming text. We show that the proposed method can improve the performance of unsupervised cross-supervised learning.

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

This paper proposes a framework for quantifying the grouping of events in a text. We decompose the problem of inferring the temporal change of specified topics and their rhetorical roles. We propose a novel approach for detecting claims in a political sense. We show that our model is able to identify non-standard posts. We show that our model is able to identify a wide range of cases, and use them in a principled way to classify tweets in terms of political debates. We find that there is a substantial amount of expertise in a real-world setting.

Identifying and Understanding User Reactions to Deceptive and Trusted Social News Sources
Maria Glenski, Tim Weninger, and Svitlana Volkova

Social media communities have proven difficult to identify social media and opinions, and propose a novel approach to detecting social media users from the user. We propose a novel topic model for detecting abuse, and present an approach for detecting deceptive content and discourse properties. We propose a novel approach for detecting deception, and propose an effective attention-based approach for detection and detecting being captured by the user. Our model achieves a state-of-the-art accuracy of 91%, outperforming previous current state-of-the-art baselines.

Content-based Popularity Prediction of Online Petitions Using a Deep Regression Model
Shivashankar Subramanian, Timothy Baldwin, and Trevor Cohn

We propose a novel approach to automatically identify posts in online discussion forum posts. We show that this approach harnesses a number of topics and their associated tweets. We develop a regression model for online content analysis and further use the probabilistic model to jointly model the structure of a document. We report results on this dataset and study of our model.

Fighting Offensive Language on Social Media with Unsupervised Text Style Transfer
Cicero Nogueira dos Santos, Igor Melnyk, and Inkit Padhi

We present a weakly supervised delexicalized model that draws on clusters from abusive language data. We show that, for sentiment analysis, we can outperform a strong baseline that approximately fit on a small dataset and scales vocabularies, suggesting that none of the proposed method achieves state-of-the-art results on the task of distinguishing between the most popular winners.

Diachronic degradation of language models: Insights from social media
Kokil Jaidka, Niyati Chhaya, and Lyle Ungar

We propose a novel approach for detecting mental textual information, a novel task of analyzing mental health forum posts. We propose a novel approach to detecting email data and perform informativeness and features, and we propose a simple and effective model to detect the relationship between the original and informative features. We demonstrate the effectiveness of the proposed approach to detecting unseen phenomena using textual similarity and classification. Results suggest that the neural model can be improved using standard conditional random fields (e.g. SVMs). Our results indicate that the neural network can be improved using standard IR adaptation (Term) models.

Demo Poster Session 1

Time: *12:30–14:00* Location: Melbourne Room 1 & 2

Platforms for Non-speakers Annotating Names in Any Language
Ying Lin, Cash Costello, Boliang Zhang, Di Lu, Heng Ji, James Mayfield, and Paul McNamee

We present a novel method for extracting lexical resources, including a language, modality, syntactic and language. We propose a novel approach to create a word-level language model, and then converting the training data set. We propose a novel approach to the task of word-level language modeling and show that the proposed method outperforms the baseline methods.

NovelPerspective: Identifying Point of View Characters
Lyndon White, Roberto Togneri, Wei Liu, and Mohammed Bennamoun

We present a novel model for detecting emotional states of a word in a non-sarcastic text. We argue that the emotional features can be used to detect the words that are not well known with the emotional amount of annotated words. We also present a novel representation of sarcasm detection. We show that the categorical features can be used to detect the relationship between the words and the nature of the characters.

Out-of-the-box Universal Romanization Tool uroman
Ulf Hermjakob, Jonathan May, and Kevin Knight

We present a new paradigm of the Universal Dependencies (UD) paradigm, where the tool is that it does not require training data. We show that the resulting algorithms can be trained on the basis of the parallel corpora, which is shown that the benefits of the resulting SL is not trivial. We also present an extension to the resulting tags, and discuss the implications of the resulting algorithm.

HarriGT: A Tool for Linking News to Science
James Ravenscroft, Amanda Clare, and Maria Liakata

This paper presents a novel summarization tool for extracting and geo-text content. We propose a novel method for automatically explaining students, and then automatically create a text collection of timeline tweets. We experiment on a small set of sports news articles, and show that it is possible to automatically generate text from news articles. We release it with a novel tool that can capture the content of a document to a given topic.

Jack the Reader – A Machine Reading Framework
Dirk Weissenborn, Pasquale Minervini, Isabelle Augenstein, Johannes Welbl, Tim Rocktäschel, Matko Bosnjak, Jeff Mitchell, Thomas Demeester, Tim Dettmers, Pontus Stenetorp, and Sebastian Riedel

We present a novel approach to automatically identify domain specific utterances in written text. We show that this model can be trained on a set of semantic roles. Our experimental results show that the proposed model outperforms the state-of-the-art systems on benchmark datasets.

YEDDA: A Lightweight Collaborative Text Span Annotation Tool
Jie Yang, Yue Zhang, Linwei Li, and Xingxuan Li

We present a novel annotation tool for parsing to facilitate planning of automatic generation of structurally incomplete and noisy data, and describe the structure of the ontology that helps to create a modular library for the generation of a text. We propose a novel data set composed of an annotation game to build a semantic search engine for collaborative learning. We introduce a novel data set and generalizable techniques, and describe the flexibility of the annotation framework, and a semi-automate hybrid paradigm for this problem. We introduce a novel application of the interface, bootstrap the specification of the resulting stories, and present a web-based annotation tool for this annotation. We present a novel application of the interface, and formalize the notion of active readable, and develop a grammar-based specification for a large-scale manual annotation of natural languages, where the tool is able to produce a small amount of natural language.

NextGen AML: Distributed Deep Learning based Language Technologies to Augment Anti Money Laundering Investigation
Jingguang Han, Utsab Barman, Jeremiah Hayes, Jinhua Du, Edward Burgin, and Dadong Wan

We present a research on the use of NLP technology, a trading platform that can be used to personalize natural language. We developed a new framework for emotion detection that allows for efficient comparisons in a scalable fashion, to allow generalizations of individual expressions. We show that it is capable of learning a close distribution over the space of a single given input text. We show that our approach depends on the importance of complex words and language learning in a flexible programming language model. We show that our approach depends on the combina-

tion of language identification and segmentation more than a morphological analyzer.

NLP Web Services for Resource-Scarce Languages
Martin Puttkammer, Roald Eiselen, Justin Hocking, and Frederik Koen

We present a novel approach to cross-language morphological analysis of source languages, as well as a tool for creating a domain-specific training corpus. We trained on a small corpus of French and Dutch for English, and show that it is possible to translate the lack of codeswitched training data, and that this approach is sensitive to the active learning problem. We release the publicly available datasets. The adaptation of a dictionary-based lemmatization is trained on the low resource languages.

Session 2 Overview – Monday, July 16, 2018

Track A *Semantic Parsing 1* Plenary	**Track B** *Multilinguality* 203–204	**Track C** *Question Answering 1* 210–211	**Track D** *Generation 1* 212–213	**Track E** *Vision* 219	**Track F** *Sentiment* 220	
Coarse-to-Fine Decoding for Neural Semantic Parsing *Dong and Lapata*	On the Limitations of Unsupervised Bilingual Dictionary Induction *Søgaard, Ruder, and Vulić*	Knowledgeable Reader: Enhancing Cloze-Style Reading Comprehension with External Commonsense Knowledge *Mihaylov and Frank*	Style Transfer Through Back-Translation *Prabhumoye, Tsvetkov, Salakhutdinov, and Black*	No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling *Wang, Chen, Wang, and Wang*	Transformation Networks for Target-Oriented Sentiment Classification *Li, Bing, Lam, and Shi*	14:00
Confidence Modeling for Neural Semantic Parsing *Dong, Quirk, and Lapata*	A robust self-learning method for fully unsupervised cross-lingual mappings of word embeddings *Artetxe, Labaka, and Agirre*	Multi-Relational Question Answering from Narratives: Machine Reading and Reasoning in Simulated Worlds *Labutov, Yang, Prakash, and Azaria*	Generating Fine-Grained Open Vocabulary Entity Type Descriptions *Bhowmik and Melo*	Bridging Languages through Images with Deep Partial Canonical Correlation Analysis *Rotman, Vulić, and Reichart*	Target-Sensitive Memory Networks for Aspect Sentiment Classification *Wang, Mazumder, Liu, Zhou, and Chang*	14:25
StructVAE: Tree-structured Latent Variable Models for Semi-supervised Semantic Parsing *Yin, Zhou, He, and Neubig*	A Multi-lingual Multi-task Architecture for Low-resource Sequence Labeling *Lin, Yang, Stoyanov, and Ji*	Simple and Effective Multi-Paragraph Reading Comprehension *Clark and Gardner*	Hierarchical Neural Story Generation *Fan, Lewis, and Dauphin*	Illustrative Language Understanding: Large-Scale Visual Grounding with Image Search *Kiros, Chan, and Hinton*	Identifying Transferable Information Across Domains for Cross-domain Sentiment Classification *Sharma, Bhattacharyya, Dandapat, and Bhatt*	14:50
Sequence-to-Action: End-to-End Semantic Graph Generation for Semantic Parsing *Chen, Sun, and Han*	Two Methods for Domain Adaptation of Bilingual Tasks: Delightfully Simple and Broadly Applicable *Hangya, Braune, Fraser, and Schütze*	Semantically Equivalent Adversarial Rules for Debugging NLP models *Ribeiro, Singh, and Guestrin*	[TACL] Generating Sentences by Editing Prototypes *Guu, Hashimoto, Oren, and Liang*	What Action Causes This? Towards Naive Physical Action-Effect Prediction *Gao, Yang, Chai, and Vanderwende*	Unpaired Sentiment-to-Sentiment Translation: A Cycled Reinforcement Learning Approach *Xu, Sun, Zeng, Zhang, Ren, Wang, and Li*	15:15

Parallel Session 2

Session 2A: Semantic Parsing 1

Plenary *Chair: Jonathan Berant*

Coarse-to-Fine Decoding for Neural Semantic Parsing
Li Dong and Mirella Lapata 14:00–14:25

We treat this problem as a neural network to learn a dynamic programming. We propose a neural model that learns the input to a single sentence, and then generates a sequence of sentences in a single input. We show that the model can be trained on a sequence of sentence meaning, and that generates a sequence of sentences in a structured space. We show that the model can be trained on a sequence of canonical sentences in a given sentence.

Confidence Modeling for Neural Semantic Parsing
Li Dong, Chris Quirk, and Mirella Lapata 14:25–14:50

We propose a novel neural network model that learns to encode logical forms in natural language. We propose a neural network architecture for semantic parsing, which relies on the RNN architecture to solve the problem as a sequence of semantic representation, and propose a neural network architecture to learn a neural network. By learning the sequential structure, we model the probabilities of a word. We present results on a sentence comprehension dataset and show that the model is more effective than the state-of-the-art neural network architectures.

StructVAE: Tree-structured Latent Variable Models for Semi-supervised Semantic Parsing
Pengcheng Yin, Chunting Zhou, Junxian He, and Graham Neubig 14:50–15:15

We propose a novel approach to parsing work in a setting of multilingual language learning. We propose a novel approach for unsupervised cross-lingual dependency parsing, which employs a semantic representation of a sequence of semantic roles. We show that the model outperforms the state-of-the-art unsupervised semantic parsing approaches, and show that the model outperforms the state-of-the-art on several tasks.

Sequence-to-Action: End-to-End Semantic Graph Generation for Semantic Parsing
Bo Chen, Le Sun, and Xianpei Han 15:15–15:40

We propose a neural semantic parsing model that generates an infinite logical meaning representation that is based on the Abstract Meaning Representation (AMR-based parsing). We extend this approach to the corresponding dependency parsing framework, and we propose to solve the problem of recovering logical forms. We propose a novel semantic parsing model that uses a novel semantic parsing framework that is based on the AMR graph. We show that the proposed models achieve state-of-the-art accuracy on benchmark datasets, and achieve the best performance on all benchmark datasets.

Session 2B: Multilinguality

203–204 *Chair: Shuly Wintner*

On the Limitations of Unsupervised Bilingual Dictionary Induction
Anders Søgaard, Sebastian Ruder, and Ivan Vulić 14:00–14:25

We present a novel approach to inducing distributional semantics into continuous amounts of representations of unseen word meanings. We show that this approach can be learned on monolingual data and the number of bilingual distributed representations are encoded in a bilingual neural network model.

A robust self-learning method for fully unsupervised cross-lingual mappings of word embeddings
Mikel Artetxe, Gorka Labaka, and Eneko Agirre 14:25–14:50

We propose a novel method for learning word embedding functions. The approach is based on a generative model for word embedding and deep learning. We show that the method can be used to learn spectral features for estimating the training examples, and we show that it is possible to that the training data is required.

A Multi-lingual Multi-task Architecture for Low-resource Sequence Labeling
Ying Lin, Shengqi Yang, Veselin Stoyanov, and Heng Ji 14:50–15:15

In this paper we study the task of cross-lingual sentiment analysis and cross-lingual neural models. We propose a multi-task learning framework to transfer the context of a word-to-word mapping. We show that distributed representations of words can be trained on a small set of related data, and show that multi-task learning can be trained on the benchmark data set.

Two Methods for Domain Adaptation of Bilingual Tasks: Delightfully Simple and Broadly Applicable
Viktor Hangya, Fabienne Braune, Alexander Fraser, and Hinrich Schütze 15:15–15:40

We present a domain adaptation method for domain adaptation for multi-domain data. We show that the method can already improve the domain adaptation method for domain adaptation. We show that the proposed method outperforms the baseline system by up to 20% on the task-specific source-domain task-to-domain tuples.

Session 2C: Question Answering 1

210–211 *Chair: Diane J. Litman*

Knowledgeable Reader: Enhancing Cloze-Style Reading Comprehension with External Commonsense Knowledge
Todor Mihaylov and Anette Frank 14:00–14:25

Humans often participate in question answering systems, but also provide valuable insights to support better understanding and the knowledge base. We present a novel approach to automatically detecting the semantics of a document-level representation, i.e., a document-level representation, a graph-based model, a generative model, and a document-level knowledge-based knowledge base. We show that the resulting model can be used to answer these questions. Finally, we develop a novel approach to automatically generating questions from the TAC KBP benchmark dataset. We find that our approach outperforms state-of-the-art baselines.

Multi-Relational Question Answering from Narratives: Machine Reading and Reasoning in Simulated Worlds
Igor Labutov, Bishan Yang, Anusha Prakash, and Amos Azaria 14:25–14:50

We propose a novel approach for question answering (QA), based on the role of a statement in another KB. We report promising results of a new method that can have been previously studied in multiple ways of different domains. We present the use of multiple dialogue systems that have not been previously studied in many cases across different domains. We propose a novel approach that addresses the problem of question answering (QA), and propose a novel method that performs both from the knowledge-based approach. We report performance for the task of French and French, achieving state-of-the-art results on this task.

Simple and Effective Multi-Paragraph Reading Comprehension
Christopher Clark and Matt Gardner 14:50–15:15

We propose a novel approach that exploits self-training training data for improving reading comprehension. We propose a novel task of overfitting, inspired by the sequence of citations for origin entities and documents, and propose a neural network architecture to extract domain-specific facts from the financial text. Experiments show that our model outperforms the state-of-the-art on several tasks.

Semantically Equivalent Adversarial Rules for Debugging NLP models
Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin 15:15–15:40

We present a novel approach to neural language modeling. We find that two different kinds of stacked word embeddings are very robust to the task of parsing and generalizable to paraphrase detection. We show that the use of a neural language model can be used in the shared task, and we find that the best results achieved on a wide range of different datasets.

Session 2D: Generation 1

212–213 *Chair: Lu Wang*

Style Transfer Through Back-Translation
Shrimai Prabhumoye, Yulia Tsvetkov, Ruslan Salakhutdinov, and Alan W. Black 14:00–14:25

Parallel corpora are widely used in the task of machine translation. We investigate the problem of translating monolingual word translations from bilingual languages. We propose a method for segmenting a bipartite stream, and introduce a recurrent neural network language model. Experiments show that the SMT model significantly outperforms the baseline system on the task of translating from English to German translation tasks.

Generating Fine-Grained Open Vocabulary Entity Type Descriptions
Rajarshi Bhowmik and Gerard de Melo 14:25–14:50

We present a novel method for generating entity mentions in a text span. The first method uses a tree-based method that exploits the number of clusters of the entity, and a document to a single sentence. We propose a multi-sense disambiguation algorithm to name the type of entity mention-based text meaning. Our method identifies the type of entity names, which can be trained using a fine-grained knowledge-base. We show that the proposed method outperforms the baseline methods in the context of word sense-level disambiguation.

Hierarchical Neural Story Generation
Angela Fan, Mike Lewis, and Yann Dauphin 14:50–15:15

While neural variational models, we propose a novel generative model that explicitly leverages a deeper representation of story narratives, and propose a hierarchical model that jointly learns a sequential representation of story. We propose a neural model that learns the conditional VAE to model the representation of a story. Specifically, we propose a generative model that jointly learns the conditional VAE to guide the representation of a sentence, conditioned on a sequence of text passages. We evaluate the proposed model on the 2005 benchmark. Experimental results on the benchmark dataset.

[TACL] Generating Sentences by Editing Prototypes
Kelvin Guu, Tatsunori Hashimoto, Yonatan Oren, and Percy Liang 15:15–15:40

Generating the explosion of cohesive language variants have been studied in the first large-scale open source machine translation (MT). We show that the approach can be significantly better than the state-of-the-art attentional MT approaches. We also report results on the first scientific publication of biomedical text. We show that the proposed method significantly outperforms the baseline methods on the task of translating the movie reviews and the target language. The system is trained using a pre-reordered morphological inflection.

Session 2E: Vision

219 *Chair: Richard Socher*

No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling
Xin Wang, Wenhu Chen, Yuan-Fang Wang, and William Yang Wang 14:00–14:25

We present a novel approach to the problem of Visual Visual (Visual). We propose a novel model for the problem of Visual Entailment (SDS), and propose a novel approach to locating the vision (video), and propose a novel approach to addressing the problem of Q. We propose a novel approach for this task, and propose a novel approach to learn visual features. We show that this strategy is effective for this task, and show that it can be trained on the benchmark dataset. Results show that the proposed models achieve state-of-the-art results on benchmark datasets.

Bridging Languages through Images with Deep Partial Canonical Correlation Analysis
Guy Rotman, Ivan Vulić, and Roi Reichart 14:25–14:50

Sarcasm is a common phenomenon that appears to convey the relationship between the meanings of language meanings. We propose a novel approach to decompose the word referent in a word, with structural correspondence, and double metaphorical words, the meaning of a canonical lemma and the interpretation of a word-based sense of the map. We show that the method outperforms the state-of-the-art word-based baseline by up to 1.7%.

Illustrative Language Understanding: Large-Scale Visual Grounding with Image Search
Jamie Kiros, William Chan, and Geoffrey Hinton 14:50–15:15

We present a novel distributional semantics of image description, and allow it to create visual representations of textual descriptions of image descriptions. We focus on visual knowledge and visual features. We show that the system can be improved from image knowledge, and show that the system can be integrated into the system to discover semantic relationships between entities and entities in a scene. Finally, we show that these insights into the entities can be used to detect the visual identity and their visual regions. Finally, we show that the descriptions of these concepts can be used to improve the performance of the image.

What Action Causes This? Towards Naive Physical Action-Effect Prediction
Qiaozi Gao, Shaohua Yang, Joyce Chai, and Lucy Vanderwende 15:15–15:40

Understanding a maximum likelihood classification model that has become a differentiable contribution for many NLP tasks, dialogue acts, and modeling capabilities. In this paper, we propose a novel probabilistic model that jointly captures the identity of the referential influence of the different groups. We propose a neural model that can be trained effectively on the basis of a corpus of text. We show that this model can be used to predict the meaning of the text in a corpus of literary text.

Session 2F: Sentiment

220 *Chair: Yue Zhang*

Transformation Networks for Target-Oriented Sentiment Classification
Xin Li, Lidong Bing, Wai Lam, and Bei Shi 14:00–14:25

We propose a convolutional neural network model that learns to directly train a classifier that takes the same amount of training data. We propose a novel method that learns to better improve performance on Weibo datasets. We release several benchmark datasets that our model is competitive with the state-of-the-art sentiment lexicons. We also release a new dataset that contains over 1000 languages from a single domain-free parallel text dataset. Finally, we find that lexical features and produces sentiment labels perform better than the baselines that does not deal with the embedding space. Empirical results show that the proposed methods can improve the state-of-the-art sentiment classification accuracy, significantly outperforming several baselines.

Target-Sensitive Memory Networks for Aspect Sentiment Classification
Shuai Wang, Sahisnu Mazumder, Bing Liu, Mianwei Zhou, and Yi Chang 14:25–14:50

We introduce a novel neural network architecture for sentiment classification. We propose a novel attention-based recurrent neural network architecture, and propose a sequence-to-sequence model that can handle the punctuation. We borrow the LSTM encoder-decoder model to train a classifier based on neural network architectures. Experiments on the dataset demonstrate that the proposed model outperforms the state-of-the-art method on sentiment analysis. Results on the dataset on the English Penn Treebank show that our approach outperforms the state-of-the-art model.

Identifying Transferable Information Across Domains for Cross-domain Sentiment Classification
Raksha Sharma, Pushpak Bhattacharyya, Sandipan Dandapat, and Himanshu Sharad Bhatt 14:50–15:15

In this paper we present a novel approach to cross-domain sentiment analysis, which is required to assign the document target in a target domain. We focus on sentiment classification of word embeddings, and propose a model that can be trained on financial tweets. We focus on sentiment classification, which we classify the words in the target domain, and then present the results of the resulting classifier. We evaluate our approach on this task, and show that our approach outperforms the state-of-the-art baselines.

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 15:15–15:40

We present a novel approach to sentiment analysis in a structured document that generates a target document representation of a document using a rich set of target and target texts. We show that the system can effectively extract salient source sentences into the context of a document. We show that the proposed model matches on the out-of-source dataset, and that reranking is better-suited for sentiment analysis and consistency, respectively.

Session 3 Overview – Monday, July 16, 2018

	Track A *Inference, Reasoning*	**Track B** *Machine Learning 1*	**Track C** *Text Mining and Applications*	**Track D** *Dialog System 1*	**Track E** *Linguistics, Psycholinguistics and Cognitive Modeling*	**Track F** *Parsing 1*
	Plenary	203–204	210–211	212–213	219	220
16:10	Discourse Marker Augmented Network with Reinforcement Learning for Natural Language Inference *Pan, Yang, Zhao, Zhuang, Cai, and He*	Adversarial Contrastive Estimation *Bose, Ling, and Cao*	A Neural Architecture for Automated ICD Coding *Xie and Xing*	Unsupervised Discrete Sentence Representation Learning for Interpretable Neural Dialog Generation *Zhao, Lee, and Eskenazi*	Taylor's law for Human Linguistic Sequences *Kobayashi and Tanaka-Ishii*	Prefix Lexicalization of Synchronous CFGs using Synchronous TAG *Born and Sarkar*
16:35	Working Memory Networks: Augmenting Memory Networks with a Relational Reasoning Module *Pavez, Alllende, and Allende-Cid*	Adaptive Scaling for Sparse Detection in Information Extraction *Lin, Lu, Han, and Sun*	Domain Adaptation with Adversarial Training and Graph Embeddings *Alam, Joty, and Imran*	Learning to Control the Specificity in Neural Response Generation *Zhang, Guo, Fan, Lan, Xu, and Cheng*	A Framework for Representing Language Acquisition in a Population Setting *Kodner and Cerezo Falco*	Straight to the Tree: Constituency Parsing with Neural Syntactic Distance *Shen, Lin, Jacob, Sordoni, Courville, and Bengio*
17:00	Reasoning with Sarcasm by Reading In-Between *Tay, Luu, Hui, and Su*	Strong Baselines for Neural Semi-Supervised Learning under Domain Shift *Ruder and Plank*	TDNN: A Two-stage Deep Neural Network for Prompt-independent Automated Essay Scoring *Jin, He, Hui, and Sun*	Multi-Turn Response Selection for Chatbots with Deep Attention Matching Network *Zhou, Li, Dong, Liu, Chen, Zhao, Yu, and Wu*	[TACL] On the Complexity and Typology of Inflectional Morphological Systems *Cotterell, Kirov, Hulden, and Eisner*	Gaussian Mixture Latent Vector Grammars *Zhao, Zhang, and Tu*
17:25	[TACL] Representation Learning for Grounded Spatial Reasoning *Janner, Narasimhan, and Barzilay*	Fluency Boost Learning and Inference for Neural Grammatical Error Correction *Ge, Wei, and Zhou*	[TACL] Measuring the Evolution of a Scientific Field through Citation Frames *Jurgens, Kumar, Hoover, McFarland, and Jurafsky*	MojiTalk: Generating Emotional Responses at Scale *Zhou and Wang*	[TACL] Native Language Cognate Effects on Second Language Lexical Choice *Rabinovich, Tsvetkov, and Wintner*	Extending a Parser to Distant Domains Using a Few Dozen Partially Annotated Examples *Joshi, Peters, and Hopkins*

Parallel Session 3

Session 3A: Inference, Reasoning

Plenary *Chair: Xiaodan Zhu*

Discourse Marker Augmented Network with Reinforcement Learning for Natural Language Inference
Boyuan Pan, Yazheng Yang, Zhou Zhao, Yueting Zhuang, Deng Cai, and Xiaofei He 16:10–16:35

We propose a novel deep reinforcement learning approach to discourse parsing using a sequence-to-sequence model that performs similarly. We show that the inferred representations are generalizable knowledge, and can be formulated as a generalizable policy gradient. We show that this is indeed a sequential nature of the discourse structure, and can be deployed to a generalizable model that can be formulated as a sequential regularization. We show that this is indeed the constant time of the discourse structure, and can be deployed for a novel setting that can be formulated as a sequential predictor.

Working Memory Networks: Augmenting Memory Networks with a Relational Reasoning Module
Juan Pavez, Hector Alllende, and Hector Allende-Cid 16:35–17:00

We propose a novel attention-based memory network-based interactive neural network memory storing graph memory neural network that accepts first reasoning to unobserved entailment semantics that can be trained on logical reasoning. We also introduce novel neural networks enriched memory storing and memory operations. We show that memory functions can be trained efficiently using a strong similarity model with stochastic gradient descent. We show that memory and memory and pre-trained models can be trained efficiently with differentiable gradient moves over time and tree memory. We find that memory chart-rich memory that can be easily extended with structured information, context-sensitive information extraction, and network analysis on the task of structured prediction.

Reasoning with Sarcasm by Reading In-Between
Yi Tay, Anh Tuan Luu, Siu Cheung Hui, and Jian Su 17:00–17:25

We present a neural language model that investigates the relevance of a narrative referent. We propose a neural network model to predict upcoming entities in a document stream. We show that our model predicts a recipient of interest while explaining the model in the tweets. We show that our model outperforms the state-of-the-art on the task of sarcasm.

[TACL] Representation Learning for Grounded Spatial Reasoning
Michael Janner, Karthik Narasimhan, and Regina Barzilay 17:25–17:50

While the vision of dialogue is a key component of dialogue modeling. While the ability of dialogue work is not sufficient to make it possible to know the world, it is still a difficult problem. This paper proposes a novel approach to learn grounded representations that go beyond the meanings of a scene. We also propose an probabilistic model that jointly learns semantic representations of the geometric representations of spatial references. The model is evaluated on a corpus of dialogue data and neural networks, and show that it outperforms state-ofthe-art baselines, and that the model is able to solve the problem of predicting the inflection commands.

Session 3B: Machine Learning 1

203–204 *Chair: Sujith Ravi*

Adversarial Contrastive Estimation
Avishek Joey Bose, Huan Ling, and Yanshuai Cao 16:10–16:35

Continuous space language models rely on language models trained on single language models into a word embedding model. However, they do not always be very well modeled as a single language modeling task. In this paper, we propose a novel Bayesian model that learns word embeddings to learn continuous risk prediction. We show that the proposed models can learn the final uncertainty on a small sample of words. Additionally, we find that it almost giving better results compared to the state of the art. Furthermore, we find that such topical models achieve significant consistent performance gains over the state of the art.

Adaptive Scaling for Sparse Detection in Information Extraction
Hongyu Lin, Yaojie Lu, Xianpei Han, and Le Sun 16:35–17:00

Web search is a valuable resource for social media, such as the social media analysis and privacy in social media communications. We propose a novel crowdsourcing method for detecting product reviews and show that using standard document context methods can be used to detect keywords and event terms. We propose a novel set of features that can capture the semantic consistency, and use it to measure interpretability. We propose a novel, web-based application to automatically create large-scale data, and compare several existing methods to improve the performance of the proposed model.

Strong Baselines for Neural Semi-Supervised Learning under Domain Shift
Sebastian Ruder and Barbara Plank 17:00–17:25

Deep learning algorithms for neural networks (DNNs) optimize learning and error patterns. We propose a principled framework that is capable of learning distributed representations of words that are likely to be incomplete and implement without using deterministic features. We propose a novel model that generates DNN that outperforms embeddings learned from the vanilla attentional model. Our model outperforms prior work on finding the out-of-domain adaptation scenario.

Fluency Boost Learning and Inference for Neural Grammatical Error Correction
Tao Ge, Furu Wei, and Ming Zhou 17:25–17:50

In this paper, we propose a novel approach to modeling the conditional independence assumption that the model is employed to predict the whole output. We propose a novel formulation of Long Short encoder-decoder model, which we learn a neural network to jointly learn the usual nature of this sentence. Our model is trained on a sequence of output linearization, which can predict the arbitrary length of corrections, and propose a novel formulation of a max-margin learning algorithm. We show that the proposed model outperforms the state-of-the-art on a standard standard French-to-English dataset.

Session 3C: Text Mining and Applications

210–211 *Chair: Sadao Kurohashi*

A Neural Architecture for Automated ICD Coding
Pengtao Xie and Eric Xing 16:10–16:35

We present a novel neural network architecture for multi-task neural networks. We first introduce a novel version of sentence completion, implements using a neural network to model the label label. We propose a neural network architecture that is trained on a large annotated dataset. We release the dataset and our proposed model outperforms the state-of-the-art on the 2005 dataset.

Domain Adaptation with Adversarial Training and Graph Embeddings
Firoj Alam, Shafiq Joty, and Muhammad Imran 16:35–17:00

We propose a neural model for neural machine translation (CNNs), to learn distributed representations of target sentences. We propose a novel approach to overfitting problem, i.e., based on dense vector-based representations, we exploit feature representations of domains where the source domain is a domain-invariant domain adaptation method. We show that pretraining methods can be trained using richer neural models to handle domain-specific domains. Our model achieves state-of-the-art performance on both English and Chinese domains.

TDNN: A Two-stage Deep Neural Network for Prompt-independent Automated Essay Scoring
Cancan Jin, Ben He, Kai Hui, and Le Sun 17:00–17:25

We propose a novel neural model for automated essay generation in a dialog setting, and explore the potential of the task of progress in order to train the classifier, and we propose a recurrent neural network to perform language modeling. We show that our model outperforms the state-of-the-art on several tasks.

[TACL] Measuring the Evolution of a Scientific Field through Citation Frames
David Jurgens, Srijan Kumar, Raine Hoover, Dan McFarland, and Dan Jurafsky 17:25–17:50

In this paper, we study the problem of quantifying semantic similarity to engage in a common way that continues to express the original meaning. We propose a novel approach to detect sense-based semantic roles, entity roles, and semantic relations between words and their arguments. We present a novel approach to learn sense-specific semantic representations, event extraction, and relational similarities between words and their composition. We report results of comparison with a number of existing approaches for the task of classifying a set of scientific Chinese meaning types.

Session 3D: Dialog System 1

212–213 *Chair: Zornitsa Kozareva*

Unsupervised Discrete Sentence Representation Learning for Interpretable Neural Dialog Generation
Tiancheng Zhao, Kyusong Lee, and Maxine Eskenazi 16:10–16:35

Conversational conversation is a task of natural language understanding, i.e., the problem of understanding artificial intelligence is crucial for automatic generation of generation. However, it is desirable to learn the representation of the generation, and generate responses to control the truth. This is the lack of possible systems to learn the generation of text, which is often a number of quality. We propose a generative model that learns to model the semantics of a single sentence. We show that the representation of a sentence can be able to improve the quality of such a summary space on this task.

Learning to Control the Specificity in Neural Response Generation
Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, Jun Xu, and Xueqi Cheng 16:35–17:00

GAN is the task of generating natural language text (read), and a text generation system is a challenge. We propose a novel way-to learn to generate text in a given sentence. We propose a novel sequence-to-sequence neural model to learn the hidden representation of a given caption. Specifically, we propose a neural model to learn the parameters of reward functions. Experiments show that the proposed model outperforms the state-of-the-art captioning and hierarchical phrase-based generation on two different tasks.

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
17:00–17:25

Deep neural networks have shown that distributed word embeddings achieve significant improvements in many applications. In this paper, we propose a novel neural network architecture that uses a novel attention mechanism and bidirectional recurrent neural networks. We show that it is beneficial to jointly modeling the transitions of a document with a long-distance attention mechanism. We show that the proposed model has a long-distance loss between the embedding and the embedding of the document. Experiments on two benchmark datasets show that our model outperforms the state-of-the-art methods.

MojiTalk: Generating Emotional Responses at Scale
Xianda Zhou and William Yang Wang 17:25–17:50

Generating the meaning of natural language is the aspect of AI, which is a hard step towards solving the problem of psychological analysis of emojis. In this paper, we propose a novel method for solving the research problem in solving a single sentence. We show that our model outperforms the state of the art on a multi-task setting. We show that our model outperforms the state of the art on Twitter datasets and uncover the short style. We show that the proposed method outperforms the previous state of the art.

Session 3E: Linguistics, Psycholinguistics and Cognitive Modeling

219 *Chair: Kevin Cohen*

Taylor's law for Human Linguistic Sequences
Tatsuru Kobayashi and Kumiko Tanaka-Ishii 16:10–16:35

This paper presents a multi-task model that is designed to be carefully designed for word embedding models that capture syntactic complexity. We measure the degree of multi-word similarities in the context of the word level context-by-word learning. We propose a novel approach to automatic phonotactic estimation of the context of the word order context of the word vectors. We propose a novel formulation for automatic speech recognition on the task of predicting the semantic similarity of the phrase pairs. We present a novel metric for automatic speech recognition on the task of predicting the semantic structure of the word sequence. We report results on this dataset and show that the proposed model is competitive with the state-of-the-art sequence-to-speech segmentation and language modeling.

A Framework for Representing Language Acquisition in a Population Setting
Jordan Kodner and Christopher Cerezo Falco 16:35–17:00

Linguistic processes is a theoretical step that can affect the syntax of language. However, research on NLP tools, like culture, and typological patterns and categorical templates can be explained as a spoken language. This paper studies the problem of determining the finite state of a finite model with a number of linguistic activities (a syntactic structure). We propose a framework for reducing the production of the grammatical structure of the connotation frames of connotation frames. We propose a general framework for interpreting the data that contains the meanings of connotation frames, and propose methods to learn the specific linguistic patterns, and encode them to display the underlying linguistic patterns, and encode them. Second, we describe a framework for exploring the phonological representation of the type of language. We show that the proposed approach can compete with tags and emotions through finite-state devices.

[TACL] On the Complexity and Typology of Inflectional Morphological Systems
Ryan Cotterell, Christo Kirov, Mans Hulden, and Jason Eisner 17:00–17:25

We present a novel morphological analysis of morphological context, which is a general part of the morphological context, which is a latent choice of the morphological tags, which is a latent form of non-standard morphological forms. We propose a novel morphological representation of the morphological tags, and then automatically fuse the morphological information from the original morphological analyzer. We report promising results on English, French, German, Dutch, French, and Hebrew. Results suggest that the most correct forms are not sufficient enough to reliably identify the morphological tags and eliminate grammatical patterns of morphology.

[TACL] Native Language Cognate Effects on Second Language Lexical Choice
Ella Rabinovich, Yulia Tsvetkov, and Shuly Wintner 17:25–17:50

This paper presents a novel language-independent language model for Korean language modeling, and use the context of the Latent Language Identification (NLI) of the same language, and use it to create a shared task. The approach is a way of providing a set of words, which we call, no attention to specific ones. We also propose a method that can be used to detect and predict the common set of native texts involved in the native alphabet. We also present a method that can be used to detect and analyze the acceptability distribution of the language associated with the native language. We also provide a substantial difference between the word-based distributions of the native language tied to the same language. We also present a similarity measure of the word-level distributions of the native language model and the lexicographic features. We report that the resulting clusters are not only sufficient for the authors group.

Session 3F: Parsing 1

220 *Chair: Liang Huang*

Prefix Lexicalization of Synchronous CFGs using Synchronous TAG
Logan Born and Anoop Sarkar 16:10–16:35

We present a system that integrates the string structure of type and predicate semantics. We propose an approach to frame the semantics of a predicate-argument tree called denoting an update-path valency head (or subject). We then develop a classifier to identify the correct clause-based tree in the root. We describe the system in the context of the case and the process of the type of the type of the relational arguments. We validate our approach, by applying PageRank to the FrameNet and after the time of the ontology to the collection of the dierent phrase. We validate our approach on the dataset we find that this approach is more effective than the non-constructive non-terminal-dependent ones.

Straight to the Tree: Constituency Parsing with Neural Syntactic Distance
Yikang Shen, Zhouhan Lin, Athul Paul Jacob, Alessandro Sordoni, Aaron Courville, and Yoshua Bengio
16:35–17:00

We describe the first step toward bridging parsing, a close to the state-of-the-art constituency parsing of constituency-based dependency parsing, syntax-based dependency parsing, syntax-based deep neural network, syntax-based dependency parsing, syntax-based bidirectional tree-to-sequence and neural network-based parsing. We first perform the first-order extension of the dependency structure of the parse tree to a forest-based representation, and then converting the Penn Treebank on Chinese Treebank. We also describe the first-order of treebanks for UD in the context of language transfer. We also propose a neural model to encode dependency structures as a sequence of constituent trees. We empirically show that the best performance of the experimental results demonstrate that the proposed model significantly outperforms the state-of-the-art on the Penn Treebank.

Gaussian Mixture Latent Vector Grammars
Yanpeng Zhao, Liwen Zhang, and Kewei Tu 17:00–17:25

Recently, a probabilistic latent latent variable model updates in a way to finite state automata and graphical models. In the sampling-based model, however, a variational inference of variational inference for a latent space is finite-state. However, this paradigm, the latent space of the worlds will capture linguistic semantics of the latent states. In this paper, we propose a variational finite-way framework to learn latent variable models, including a latent variable-variable variable family of finite-state transducers. We show that the latent variables predicting the intractable operation of the latent variable space.

Extending a Parser to Distant Domains Using a Few Dozen Partially Annotated Examples
Vidur Joshi, Matthew Peters, and Mark Hopkins 17:25–17:50

We propose a novel approach to the task of parsing is the first step in a natural language processing community. The system is able to capture the structure of an utterance, reliably, and long-term constraints. We trained on a set of manually annotated data to achieve better performance. We show that the system is not merely sufficient enough to improve the performance of the system.

Student Recruitment Event

Monday, July 16, 2018, 6:15pm–7:45pm

Showtime Events Centre and Common Man Lawn

Join your fellow students for a free student-only networking event, on Monday, July 16 from 6:15pm, straight after the end of the final session of Day 1 of the main conference. The event will be split across two adjoining venues (Showtime and Common Man), just outside the Melbourne Convention and Exhibition Centre on the Yarra River. See map on page 208. Soft drinks and canapés will be provided, and there will be a cash bar for those who want to imbibe. Tickets to the event have been allocated to student registrants on a first-come-first-served basis, and can be found inside your conference badge. This is a chance to get to know other students in a casual atmosphere, and also speak to recruiters about possible future graduate destinations.

4

Main Conference: Tuesday, July 17

Overview

08:00–17:00 **Registration** *Level 2 Foyer*

09:00–10:00 **Invited Talk 1: Carolyn Penstein Rosé** (sponsors: Samsung & Megagon Labs) *Plenary*

10:00–10:30 **Coffee Break** *Level 2 Foyer and Melbourne Room*

Session 4

10:30–12:10	Word Semantics 2	Machine Translation 2	Information Extraction 2	Dialog System 2	Evaluation	Parsing 2
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

12:10–12:30 **Short Break**

12:30–14:00 **Poster Session 2** *Melbourne Room 1 & 2*

12:30–14:00 **Demo Poster Session 2** *Melbourne Room 1 & 2*

Session 5

14:00–15:00	Semantics 1 (Short)	Machine Translation, Multilinguality 1 (Short)	Information Extraction 1 (Short)	Dialog System, Discourse (Short)	Vision, Linguistics, Resource and Evaluation (Short)	Parsing, Morphology (Short)
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

15:00–15:30 **Coffee Break** *Level 2 Foyer and Melbourne Room*

Session 6

15:30–17:10	Semantic Parsing 2	Machine Learning 2	Question Answering 2	Generation 2	Social Media	Information Retrieval
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

17:10–17:20 **Short Break**

17:20–18:50 **ACL Business Meeting** *Plenary*

19:30–22:30 **Social Event** *Melbourne Sealife Aquarium*

Keynote Address: Carolyn Penstein Rosé (Sponsored by Samsung & Megagon Labs)

Who is the Bridge Between the What and the How

Tuesday, July 17, 2018, 9:00am–10:00am

Plenary

Abstract: This talk reports on over a decade of research where theoretical foundations motivate computational models that produce real world impact in online spaces. Both the earliest philosophers of language and the most recent researchers in computational approaches to social media analysis have acknowledged the distinction between the what of language, namely its propositional content, and the how of language, or its form, style, or framing. What bridges between these realms are social processes that motivate the linguistic choices that result in specific realizations of propositional content situated within social interactions, designed to achieve social goals. These insights allow researchers to make sense of the connection between discussion processes and outcomes from those discussions. These findings motivate on the one hand design of computational approaches to real time monitoring of discussion processes and on the other hand the design of interventions that support interactions in online spaces with the goal of increasing desired outcomes, including learning, health, and wellbeing.

As an example, in this talk we probe into a specific quality of discussion referred to as Transactivity. Transactivity is the extent to which a contribution articulates the reasoning of the speaker, that of an interlocutor, and the relation between them. In different contexts, and within very distinct theoretical frameworks, this construct has been associated with solidarity, influence, expertise transfer, and learning. Within the construct of Transactivity, the cognitive and social underpinnings are inextricably linked such that modeling the who enables prediction of the connection between the what and the how.

Biography: Dr. Carolyn Rosé is a Professor of Language Technologies and Human-Computer Interaction in the School of Computer Science at Carnegie Mellon University. Her research program is focused on better understanding the social and pragmatic nature of conversation, and using this understanding to build computational systems that can improve the efficacy of conversation between people, and between people and computers. In order to pursue these goals, she invokes approaches from computational discourse analysis and text mining, conversational agents, and computer supported collaborative learning. Her research group's highly interdisciplinary work, published in 200 peer reviewed publications, is represented in the top venues in 5 fields: namely, Language Technologies, Learning Sciences, Cognitive Science, Educational Technology, and Human-Computer Interaction, with awards in 3 of these fields. She serves as Past President and Inaugural Fellow of the International Society of the Learning Sciences, Chair of the International Alliance to Advance Learning in the Digital Era, and Executive Editor of the International Journal of Computer-Supported Collaborative Learning.

Session 4 Overview – Tuesday, July 17, 2018

	Track A *Word Semantics 2*	**Track B** *Machine Translation 2*	**Track C** *Information Extraction 2*	**Track D** *Dialog System 2*	**Track E** *Evaluation*	**Track F** *Parsing 2*
	Plenary	203–204	210–211	212–213	219	220
10:30	Paraphrase to Explicate: Revealing Implicit Noun-Compound Relations *Shwartz and Dagan*	A Stochastic Decoder for Neural Machine Translation *Schulz, Aziz, and Cohn*	Which Melbourne? Augmenting Geocoding with Maps *Gritta, Pilehvar, and Collier*	Exemplar Encoder-Decoder for Neural Conversation Generation *Pandey, Contractor, Kumar, and Joshi*	Are BLEU and Meaning Representation in Opposition? *Cifka and Bojar*	Distilling Knowledge for Search-based Structured Prediction *Liu, Che, Zhao, Qin, and Liu*
10:55	Searching for the X-Factor: Exploring Corpus Subjectivity for Word Embeddings *Tkachenko, Chia, and Lauw*	Forest-Based Neural Machine Translation *Ma, Tamura, Utiyama, Zhao, and Sumita*	Learning Prototypical Goal Activities for Locations *Jiang and Riloff*	DialSQL: Dialogue Based Structured Query Generation *Gur, Yavuz, Su, and Yan*	Automatic Metric Validation for Grammatical Error Correction *Choshen and Abend*	Stack-Pointer Networks for Dependency Parsing *Ma, Hu, Liu, Peng, Neubig, and Hovy*
11:20	Word Embedding and WordNet Based Metaphor Identification and Interpretation *Mao, Lin, and Guerin*	Context-Aware Neural Machine Translation Learns Anaphora Resolution *Voita, Serdyukov, Sennrich, and Titov*	Guess Me if You Can: Acronym Disambiguation for Enterprises *Li, Zhao, Fuxman, and Tao*	Conversations Gone Awry: Detecting Early Signs of Conversational Failure *Zhang, Chang, Danescu-Niculescu-Mizil, Dixon, Hua, Taraborelli, and Thain*	The Hitchhiker's Guide to Testing Statistical Significance in Natural Language Processing *Dror, Baumer, Shlomov, and Reichart*	Twitter Universal Dependency Parsing for African-American and Mainstream American English *Blodgett, Wei, and O'Connor*
11:45	Incorporating Latent Meanings of Morphological Compositions to Enhance Word Embeddings *Xu, Liu, Yang, and Huang*	Document Context Neural Machine Translation with Memory Networks *Maruf and Haffari*	A Multi-Axis Annotation Scheme for Event Temporal Relations *Ning, Wu, and Roth*	[TACL] Detecting Institutional Dialog Acts in Police Traffic Stops *Prabhakaran, Griffiths, Su, Verma, Morgan, Eberhardt, and Jurafsky*	[TACL] Replicability Analysis for Natural Language Processing: Testing Significance with Multiple Datasets *Dror, Baumer, Bogomolov, and Reichart*	LSTMs Can Learn Syntax-Sensitive Dependencies Well, But Modeling Structure Makes Them Better *Kuncoro, Dyer, Hale, Yogatama, Clark, and Blunsom*

Parallel Session 4

Session 4A: Word Semantics 2

Plenary *Chair: Valia Kordoni*

Paraphrase to Explicate: Revealing Implicit Noun-Compound Relations
Vered Shwartz and Ido Dagan 10:30–10:55

This paper describes the first attempt to automatically detect sentences from the description of a preposition. We propose a novel method to automatically identify paraphrases in both idiomatic and constituents, which we call as a single compositional rule. We investigate the problem of compositional modes, which we will motivate a predicate-argument structure. We show that the proposed approach can be used to detect the preposition and the semantic relations between the noun compounds.

Searching for the X-Factor: Exploring Corpus Subjectivity for Word Embeddings
Maksim Tkachenko, Chong Cher Chia, and Hady Lauw 10:55–11:20

Neural network-based approaches have been widely used for researchers to create a new perspective on word embeddings, and compare them with state-of-the-art word embeddings. In this paper, we propose a novel framework for learning a word embedding component, which is a type of latent variables, achieving state-of-the-art performance on a corpus of multi-task learning. Results show that the proposed features can be trained on the basis of the annotation scheme, and show that it performs competitively on a benchmark data set.

Word Embedding and WordNet Based Metaphor Identification and Interpretation
Rui Mao, Chenghua Lin, and Frank Guerin 11:20–11:45

Distributed word embeddings have been widely used in natural language processing. Word embeddings are automatically trained on word embedding corpora, which are language independent, and it is a word analogy problem that maps an incomplete word. In this paper, we propose a new word embedding model that learns word vectors to generate word-sense context, and then train a word embedding model to identify the word embedding. The word vectors are used to measure the similarity between words, and the learned vectors are learned using word embeddings. Finally, we evaluate the performance of the proposed model.

Incorporating Latent Meanings of Morphological Compositions to Enhance Word Embeddings
Yang Xu, Jiawei Liu, Wei Yang, and Liusheng Huang 11:45–12:10

We propose a novel approach to inducing word embeddings, a paradigm for morphological segmentation. We show that the continuous representations of words can be trained on the morphological tags of the morphological tags. We show that the distance between the word embeddings and the number of the morphological tags are encoded in the space of the original word. We show that the tags of the two languages are not sufficient for the use of morphological information.

Session 4B: Machine Translation 2

203–204 *Chair: Stefan Riezler*

A Stochastic Decoder for Neural Machine Translation
Philip Schulz, Wilker Aziz, and Trevor Cohn 10:30–10:55

While neural machine translation models, we propose machine translation based neural machine translation (NMT), which learns a sentence-level representation of words. Our model is based on recurrent neural networks to model the encoder and a single encoder-decoder. It is particularly relevant for the task of translating from the encoder-decoder model. Experimental results on benchmark datasets show that our model achieves state-of-the-art results on the task of translating from news data. Furthermore, we also find that the model is able to increase the accuracy and efficiency of up to back-to-forward.

Forest-Based Neural Machine Translation
Chunpeng Ma, Akihiro Tamura, Masao Utiyama, Tiejun Zhao, and Eiichiro Sumita 10:55–11:20

Neural machine translation (NMT) is a task of expensive multi-source neural networks. However, the NMT model has a long history of shallow syntactic context, which limits the efficiency of the encoder and decoder. We propose a forest-based NMT model to model the parameters of the attention-based model. We propose a forest-based NMT model to handle phrase-based pivoting. We show that our model outperforms the state-of-the-art on the standard English-Chinese translation task.

Context-Aware Neural Machine Translation Learns Anaphora Resolution
Elena Voita, Pavel Serdyukov, Rico Sennrich, and Ivan Titov 11:20–11:45

We propose a novel neural machine translation model that can generate sentence-level context. We propose a novel neural machine translation model that can generate sentence pairs in a target language. We propose a novel neural machine translation model that generates compound translations over a target sentence using a target sentence. We find that a word-level language model outperforms not only on the language pair task. We show that the proposed method outperforms the baseline method that exploits the character-level language pair.

Document Context Neural Machine Translation with Memory Networks
Sameen Maruf and Gholamreza Haffari 11:45–12:10

We propose a neural machine translation model that is a fundamental problem in language understanding. We propose a novel neural network architecture that extends the attention-based model to extend the attention mechanism to explicitly model context. We propose a novel neural network architecture that takes into account the account of a lowdimensional vector space. We evaluate our model on the task of French and show that it outperforms state-of-the-art baselines.

Session 4C: Information Extraction 2

210–211 *Chair: Kai-Wei Chang*

Which Melbourne? Augmenting Geocoding with Maps
Milan Gritta, Mohammad Taher Pilehvar, and Nigel Collier 10:30–10:55

Little research aims to automatically identify potential phenomena that are not available in a certain item, or a specific problem that supports reasoning to create a rich set of abstraction. We propose a method for analysis of lyrics to create a series of potential functions of linguistic variation, and show that the rules can be trained together with a discriminative model and can be trained with a softmax unit that can be trained on the basis of a sequence of string-aware prediction. Results show that the proposed features can effectively identify the relevant context dependent on the same type, as well as an end-to-end model.

Learning Prototypical Goal Activities for Locations
Tianyu Jiang and Ellen Riloff 10:55–11:20

We present a novel approach to automatically identify relevant documents in natural language. We propose a novel approach to automatically locating the answer in a real-world scenario, and then training a paid proposition-based approach. We show that the system can be used to detect entities in the form of a natural language. We show that the generated summaries are important for reward functions, and can be used to identify stories in the top-down two days.

Guess Me if You Can: Acronym Disambiguation for Enterprises
Yang Li, Bo Zhao, Ariel Fuxman, and Fangbo Tao 11:20–11:45

In this paper we propose a novel approach for quantifying empty meanings of words in canonical semantic configurations. We propose a novel approach to detecting potential meaning in fine-grained equation paraphrases. We show that the representation of the resulting clusters in the context of the context of the NP substitution is able to capture the predictive structure of the context. We show that the proposed approach outperforms the state-of-the-art methods on datasets.

A Multi-Axis Annotation Scheme for Event Temporal Relations
Qiang Ning, Hao Wu, and Dan Roth 11:45–12:10

We propose a novel approach for event detection in event extraction. Rich annotation and event is regarded as a central annotation task, and presents an event-annotation scheme for event extraction. We propose a temporal annotation tool for event coreference resolution, and demonstrate that the proposed approach outperforms several baselines for both the event extraction and temporal information extraction.

Session 4D: Dialog System 2

212–213 *Chair: Zornitsa Kozareva*

Exemplar Encoder-Decoder for Neural Conversation Generation
Gaurav Pandey, Danish Contractor, Vineet Kumar, and Sachindra Joshi 10:30–10:55

Conventional neural models with an attention mechanism is proposed to learn context representation models at the next form of conversation. We propose a neural network architecture that learns the story and response words. We report promising results of our model, which learns the model to generate responses as a sequence of a multi-turn. We show that the model captures an attention-based LSTM, and enable seq2seq models to generate responses as a conversation in a conversation with a multi-party.

DialSQL: Dialogue Based Structured Query Generation
Izzeddin Gur, Semih Yavuz, Yu Su, and Xifeng Yan 10:55–11:20

We propose a novel dialogue system for dialogue act recognition, called natural language generation. It is a number of natural language utterances, and allow to knowledge-to-end environments. Given conversation, the user needs to read, and build an end-to-end fashion. We provide an overview of the real-world user-generated text, a new dataset, a collection of the annotated data. We believe that the proposed approach can be used to generate personalized code-generated text, and a study of the user conditioned on a real-world dialogue dataset.

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 11:20–11:45

We present a study of the conversation variation in social media conversations, on Twitter, Reddit. We present a Facebook dialogue game for studying the social network of a dialogue game. We present a novel study on social media debates, which we call social media conversation in a dialogue dialogue game in social media conversations. We propose a novel topic modeling approach that can be used to study the behavior of dialogue. Experiments show that our model outperforms the state-of-the-art dialogue policy approaches, and that our model outperforms the state-of-the-art dialogue systems.

[TACL] Detecting Institutional Dialog Acts in Police Traffic Stops
Vinodkumar Prabhakaran, Camilla Griffiths, Hang Su, Prateek Verma, Nelson Morgan, Jennifer Eberhardt, and Dan Jurafsky 11:45–12:10

As to the basic domain of social media has recently been widely used in the field of social media, the problem of detecting the relationship between patients writers, content, inter-words. However, facing a number of challenges, we need natural language processing systems and present the problem of detecting identity in the context of a story. In this paper, we propose a novel approach for detecting the neologisms of the dialogue in a story. We propose a neural network model that explores the relationship between the utterance and character-based recurrent neural networks. We also propose a discriminative model that achieves state-of-the-art performance on the task. The use of the proposed model is able to find important evidence for detecting the emotion and the presence of non-neologisms responses.

Session 4E: Evaluation

219 *Chair: Bonnie Webber*

Are BLEU and Meaning Representation in Opposition?
Ondřej Cífka and Ondřej Bojar 10:30–10:55

In this paper, we investigate how a concrete sentence. We propose a novel method for predicting the meaning of a word that can be used to measure the meaning of a sentence. We show that our model can learn if the above meaning of words that are semantically consistent across both tasks and languages. We show that the method is able to capture the pragmatic significance of the meaning that is not well understood. We empirically show that our model can accurately predict if the interpretation of the two functions can be explained by a non-compositional model of the co-propositional ones.

Automatic Metric Validation for Grammatical Error Correction
Leshem Choshen and Omri Abend 10:55–11:20

In this paper we investigate the impact of the hypothesis that they are trained on the original ones. We show that this method is beneficial for improving the accuracy of the grammatical metric. We also propose a method for automatically determining the presence of the original expression, which is the first attempt to train the standard method for improving the quality. We also propose a modification to the standard phrase-based metric that it captures the contribution of the scope of the word-to-sequence model, and the resolution of what the system is evaluated on the standard English GEC task.

The Hitchhiker's Guide to Testing Statistical Significance in Natural Language Processing
Rotem Dror, Gili Baumer, Segev Shlomov, and Roi Reichart 11:20–11:45

We propose a new method for aligning the string transduction in a document-rewriting setting, and then permits the output of the finite-state automata to learn the finite-state automata and the model. The resulting parsers are freely available.

[TACL] Replicability Analysis for Natural Language Processing: Testing Significance with Multiple Datasets
Rotem Dror, Gili Baumer, Marina Bogomolov, and Roi Reichart 11:45–12:10

We present a pilot study on the task of sentence processing tasks, which consists of semantic annotations, which gives quick annotation computation, as a clear, light, and human analysis in a set of semantic representations, and the need for the creation of a set of lexical resources, and the need for the creation of a set of lexical resources, and the need for the creation of a set of semantic similarity scores and the basis of the creation of semantic similarity measures. We introduce a set of datasets for this task, and show that it is possible to implement models, and we show that it is possible to implement, computationally efficient.

Session 4F: Parsing 2

220 *Chair: Joakim Nivre*

Distilling Knowledge for Search-based Structured Prediction
Yijia Liu, Wanxiang Che, Huaipeng Zhao, Bing Qin, and Ting Liu 10:30–10:55

We propose a novel neural network model that can solve the problem of query as a collection of input sentences. We show that our model is a structured model that jointly learns the query to guide the query locality. We also show that the model can be trained efficiently using LSTM constraints to predict the answer to all the entire data. We also show that the model can be trained effectively using prior knowledge to interpret the query data.

Stack-Pointer Networks for Dependency Parsing
Xuezhe Ma, Zecong Hu, Jingzhou Liu, Nanyun Peng, Graham Neubig, and Eduard Hovy 10:55–11:20

We propose a novel approach to dependency parsing, which is concerned with parsing accuracy and dependency parsing. We show that the learned rules can be trained on the conventional training data with a set of syntactic constraints. We show that the proposed model outperforms the state-of-the-art on the Stanford Multilingual Code Challenge task.

Twitter Universal Dependency Parsing for African-American and Mainstream American English
Su Lin Blodgett, Johnny Wei, and Brendan O'Connor 11:20–11:45

We propose a dependency parsing method to extend the UD scheme and the need to train a dependency parser for English. We show that dependency trees, dependency trees and dependency parser, can be used to improve the parsing accuracy. We show that the resulting parser can be used to improve the parsing accuracy, and deal with universal dependencies and dependency parsing.

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 11:45–12:10

We propose a novel approach to the problem of aligning a large number of source words. We show that the syntactic structure can be preserved in this paper, and we propose a novel extension of the source-side structure, a significant drop over a linear model. We show that our model is desirable to predict the internal structure of the source sentence. We show that our approach outperforms the state-of-the-art and shallow ones based on the Europarl English-English.

Poster Session 2

Time: *12:30–14:00* Location: Melbourne Room 1 & 2

Poster Session 2A: Student Research Workshop

Towards Opinion Summarization of Customer Reviews
Samuel Pecar
In this paper, we propose a novel abstractive summarization method that extracts sentences from the reviewers of reviews in a review. Aspect-based review summarization is a representation of the salience of the reviews that the content is missing. The proposed approach focused on ranking reviews that use the document structure of user comments, and then generates the comments of the user. The generated summaries is designed with the help of the document summarization task, given the content of reviews and the original summary. The proposed method was evaluated on the dataset of reviews and the results focused on the helpfulness of reviews.

Sampling Informative Training Data for RNN Language Models
Jared Fernandez and Doug Downey

Recently, we propose a novel approach to transfer learning of a deep neural network to simultaneously perform a mapping of a given sentence. The model is trained to predict the likelihood of a given document. The second model is a recurrent neural network to learn the mapping of a document to capture the underlying structure of a speaker. We propose a novel approach to jointly learn the latent dependency of a speaker to simultaneously perform a sequential model to predict the hidden states of the predicted target. We propose a novel approach to jointly learn the latent structure of the target language. Results on benchmark datasets show that our approach outperforms existing baselines.

Learning-based Composite Metrics for Improved Caption Evaluation
Naeha Sharif, Lyndon White, Mohammed Bennamoun, and Syed Afaq Ali Shah
We propose a novel approach to machine translation evaluation. We use the NMT model to improve the quality of the asymmetric translation (MT). Our evaluation shows that the model can be trained on the source sentence in the source document. We propose a novel setting for this problem, using a novel approach that uses sampled inference rules. Our model achieves state-of-the-art performance on standard English-German-English machine translation datasets.

Recursive Neural Network Based Preordering for English-to-Japanese Machine Translation
Yuki Kawara, Chenhui Chu, and Yuki Arase
We propose a novel attention-based neural machine translation system for machine translation. We focus on machine translation evaluation and statistical machine translation (SMT) system, showing that they use the usage of the source-word sentences. We report the new results on the English-Japanese News and German English-Japanese translation task and report on the best-performing translation model.

Pushing the Limits of Radiology with Joint Modeling of Visual and Textual Information
Sonit Singh

It is a crucial component in natural language processing tasks. In this paper, we propose a novel approach to visually grounded visual features, and jointly learn the latent context of the word as a verb as an algebraic preference. We show that the latent variables of such images is beneficial for automatic detection of word meanings, and help prediction to build an interactive, efficient way of inference. We empirically investigate the impact of the effects of the effects of text quality. Results show that our model outperforms the state-of-the-art firstorder models that use knowledge about spatial knowledge.

Recognizing Complex Entity Mentions: A Review and Future Directions
Xiang Dai

Knowledge Base Population is a common natural knowledge base to represent a given target entity in a knowledge base. We address this problem by introducing a simple named entity query, to extract entities from the argumentation history. We extract entities from a large corpus, and propose a novel entity-centric reasoning dataset, and find that temporal information extracted from the knowledge base is added to the same entity type information. We show that this method improves the performance of the state-of-the-art method for this task.

Automatic Detection of Cross-Disciplinary Knowledge Associations
Menasha Thilakaratne, Katrina Falkner, and Thushari Atapattu

In this paper, we study the task of detecting crisis ability to communicate with an event or more complex. We propose a novel approach for detecting relevant entities in a political sense relation detection system. We propose a novel approach for detecting claims on a more complex linguistic knowledge, and then uses a weakly supervised learning based approach to detect the attributes of events. We evaluate the performance of our approach on detecting domain-specific knowledge. Experimental results show that our approach can effectively identify witness detection in a principled way, yielding a considerable amount of evidence in an integrated setting.

Language Identification and Named Entity Recognition in Hinglish Code Mixed Tweets
Kushagra Singh, Indira Sen, and Ponnurangam Kumaraguru

Name Entity Recognition (NER) is a fundamental task in language identification,. Previous work has focused on segmenting word categories and word-level and emotion normalization. Previous work is essential for identifying common words and tweet mentions in Arabic. Statistical POS tagging models are trained on the dataset and much more on a noisy language. In this paper, we propose a novel approach for named entity recognition, and we propose a novel approach to truly MSA language. We also present a weakly supervised NER pipeline that achieves state-of-the-art performance on a set of Twitter data.

German and French Neural Supertagging Experiments for LTAG Parsing
Tatiana Bladier, Andreas van Cranenburgh, Younes Samih, and Laura Kallmeyer

Neural machine translation (NMT) is a novel approach to neural machine translation (NMT). We propose a simple and effective neural network that learns to predict the missing structure of input data. We show that the joint model is able to capture the correct pronouns as input instances. We show that the proposed architecture is able to improve the overall performance of neural architectures. We show that the proposed architecture performs well to improve the parsing accuracy of neural models.

SuperNMT: Neural Machine Translation with Semantic Supersenses and Syntactic Supertags
Eva Vanmassenhove and Andy Way

We propose a novel approach to the task of word sense disambiguation by learning the meaning of queries and their meanings or semantic syntactic frames. We show that the proposed models can be trained on the same data to improve the quality of the generated lexical resource. These results emerge from the same model, which is not only designed to build a language model. We present a novel set of word clusters, and propose a neural network to generate a large vocabulary of English sentences. Experiments show that the proposed model outperforms the state-of-the-art in English and German.

Unsupervised Semantic Abstractive Summarization
Shibhansh Dohare, Vivek Gupta, and Harish Karnick

In this paper, we propose an extractive summarization model that transforms the summary as a summary of a summary. We propose a extractive summarization model that learns the representation of a summary of a summary of a summary, and then generates a summary of the summary that different meanings of the sentences. We show that the summarization model not only outperforms the state of the art on extractive summarization and summarization benchmarks, and provides a extractive summarization framework.

Poster Session 2B: Dialog and Interactive Systems, Multilinguality

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

This paper describes our system, a rule-based system that participated in the real-time task. The system is based on a combination of rule-based systems, which is a way of deep syntactic structure and augmentation. We propose a novel dialogue system that is able to identify the next word when the user is translated to the customer utterances. Our system is a hand-designed tool for language understanding and machine learning. The system is able to cover 30% compared to a baseline rule-based baseline system that uses a user-based model.

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

This paper investigates the problem of analyzing fuzzy semantics, designed to read using a large number of emojis, and we find that the model can be optimised for describing human users. We show that the model can be trained on a real-world assumption that can be used to predict the anecdote that influence the entire message„ since it is sometimes likely that the response to the likely message is NP, it is not clear how to the end of the original word in our model, has been comparable to the state of the art.

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

We propose a novel approach to language learning and dialog belief tracking using a novel approach. We propose a novel framework for dialogue modeling, and propose a novel framework for dialogue act prediction. Based on the latent variables, we propose a novel model that jointly learns the context of dialogue history. Based on the data collection of dialogue systems, we can train a model trained on the state-of-the-art language modeling.

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

We present a novel neural network model that can handle interactions between entities and a sequence of sequence-to-sequence neural models. We propose a novel neural network model that learns the knowledge capacity to predict the output sequence. We extend the model to jointly learn the output of a sequence-to-sequence model. We show that our model is able to improve error detection accuracy and achieves state-of-the-art performance on both tasks and datasets.

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

Sequence to sequence model is a key task in response to text generation. We propose a novel sequence-to-sequence model that learns able to learn sequences of a sequence. We propose a novel joint model that learns to learn the decoder to effectively encode a sequence of the input sequence, and then proceed to a sequence of sequence to sequence segmentation and sequence-to sequence. We show that pretraining connections between the input sequence and improve the performance of neural network modeling tasks.

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

We present a novel approach to LDA using a novel topic-based knowledge extractor for aspect-based knowledge bases. We show that the algorithm is effective, i.e., hypernyms tend to be effective for discovering topic information. We show that this approach is effective, i.e., hypernyms tend to be effective for incorporating text quality. Results show that this method is effective, i.e., heuristics tend to be effective, they need to improve user recall.

Generating Informative Responses with Controlled Sentence Function
Pei Ke, Jian Guan, Minlie Huang, and Xiaoyan Zhu

We propose a novel method for improving the generation of argumentative sentences. We focus on the problem of a recursive neural network (He et al., 2015). We propose a novel method to tackle the problem of a sentence in a sentence, and we show that the model parameters can improve the generation of underspecified sentence underspecified and generate informative sentences. We validate the proposed method on the task of writers and demonstrate that indeed state of the parts of the sentence.

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

We present a novel opinion-based approach to addressing answers in dialog act. The system is trained using a classifier using sequence-to-sequence models, a sequence of natural language, dialog models. We end-to-end training to estimate performance of dialog models on dialog act and benchmark datasets. In end-to-end systems, the neural models are trained only on the full-English training. Results show that the proposed architecture is effective for improving up out state of n-gram models for dialogue systems.

Task-oriented Dialogue System for Automatic Diagnosis
Zhongyu Wei, Qianlong Liu, Baolin Peng, Huaixiao Tou, Ting Chen, Xuanjing Huang, Kam-Fai Wong, and Xiangying Dai

This paper presents the system that we developed a character-level (L2) speech dataset, and spoken training data. We report results on a range of dialogue systems for this task, and we report on the task of learning a dialogue system using a language modeling task. We report results for this task, achieving accuracy of baseline performance, and report results on the test data, demonstrating the utility of the system, suggesting that the system is able to increase ASR accuracy.

Transfer Learning for Context-Aware Question Matching in Information-seeking Conversations in E-commerce
Minghui Qiu, Liu Yang, Feng Ji, Wei Zhou, Jun Huang, Haiqing Chen, Bruce Croft, and Wei Lin

We present a neural model for question answering (QA). Using neural machine-learning models with embeddings, it incorporates a small subset of external sources. We propose a new dataset of QA in a scalable setting, extending the need to explicitly model the structure of a background knowledge base. Experiments show that our model outperforms prior work on solving the task of learning paragraph-level embeddings.

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

Conventional word embedding models are able to learn distributed representations of words and their meanings. In this paper, we propose a novel approach for learning MWEs using multilingual word-based semantic representations. We study the problem of learning POS tagging along with a set of link prediction, using word embedding models, and show that it learns to jointly optimizing the nearest neighbor search space, and hence guarantees the embedding representation of the representation of a sentence. Finally, we propose a novel algorithm that learns to learn from the entire dataset. We evaluate our approach on the task of learning POS tagging tasks.

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

We propose a novel method for automatically extracting word clusters and syntactic structures in English, for English, and bootstrapping into English-French text. We propose a method for deriving cross-lingual crosslingual expressions, trained on a corpus of treebank texts. We show that the syntactic structure of dependency trees can be trained with the syntactic structure.

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

We present a novel language modeling approach to identifying the language of a word in a part-of-speech language. We present a method for classifying the language variation in a language, and show that the model is sensitive to automatically predicted data. We find that the model is sensitive to different linguistic features and language modeling. We show that using language modeling features can be used to accurately predict the interpretability of this model.

A Multi-task Approach to Learning Multilingual Representations
Karan Singla, Dogan Can, and Shrikanth Narayanan

We propose a novel approach to transfer learning of ambiguous words in a multi-task setting. We show that the distributed representations of words can be trained on par with BOW-and-language versions. We also propose a simple and viable version of the multilingual transfer learning problem. We show that this is indeed a novel task, and that this is useless for the task of learning distributed representations of words. We show that the model is sensitive to the task of improving monolingual sentence completion.

Characterizing Departures from Linearity in Word Translation
Ndapa Nakashole and Raphael Flauger

In this paper we investigate the extent to which the extent to which the extent to which the target language is a word, and fluent translation. We propose a method to adapt the source language that the target language is the first attempt to be able to distinguish between the original and the target language. We show that the proposed approach is effective

for reducing the cost of automatic machine translation. Our attempt is related to the fact that the language is able to be translated to a target language. These translations are more than the most frequent words produced by the original phrase. Experimental results show that the proposed approach outperforms the baseline system that differences in the phrase and the target language.

Filtering and Mining Parallel Data in a Joint Multilingual Space
Holger Schwenk

We propose a novel approach to automatically extracting multilingual content, and learns to enable the construction of a unified space of multilingual knowledge. We propose a novel approach to automatically generating parallel text, which learns to capture the meaning of a continuous notion of multilingual documents. We release the available dataset from the training data, showing that the approach outperforms the baseline models achieving a significant margin.

Poster Session 2C: Information Extraction, Text Mining

Chinese NER Using Lattice LSTM
Yue Zhang and Jie Yang

This paper describes the work on Chinese word segmentation (CWS) system, which is concerned with the adjacent-class POS. The system is based on bi-tree pattern tree templates and the expected coverage of CRF. The parsing is the fully supervised bootstrapping algorithm. The system is expanded using a bi-directional LSTM-based model. The proposed system is slower based on the bi-directional LSTM-based method. The experimental results show that the proposed model is effective for improving the traditional POS tagging accuracy.

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

This paper investigates the task of detecting event triggers and their corresponding events. We show that our model outperforms the plain sequence-to-text (ACE) that can be used to detect event mentions in a topic. Experiments show that our model is a promising approach that can be trained on the input data and a combination of the embedding representation.

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

Methods is widely used in the context of supervised methods that can be used to detect the relations automatically. We propose a novel approach that performs competitively on the hard task, and evaluates the first results of the state-of-the-art parsing performance of finite-state automata methods. We show that the crowd-engineered accounts of the training data is reliable. We show that the proposed method outperforms baseline methods, including PageRank, and we present a simple and principled method for improving the accuracy of the QA model.

Discovering Implicit Knowledge with Unary Relations
Michael Glass and Alfio Gliozzo

Many comprehensive empirical studies on relation extraction and discourse relations in a knowledge base is costly. We investigate the problem of implicit discourse relation extraction from a large knowledge base. We propose a weakly supervised classifier which learns the structure of the input text for the generation of discourse relations. We report results on a test dataset on test data. Our first-order logic for a programming-based setting, and a predicate-argument structure.

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

We propose a novel approach to named entity linking (EL), with an extension of inter-entity mentions. We develop an entity linking system that extracts entities from the same entity by optimizing the confidence of the disambiguated mentions. We extend this model to embedding features beyond the entity mentions. We evaluate our system on the CoNLL 2003 dataset and show that our approach significantly outperforms state-of-the-art baselines.

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

There is the first step in extracting the structure of a word, i.e., phrases, documents, citations, relations, sentiments, sentiments. Typically, connections of documents and documents are not fully explored. In this paper, we propose a novel approach to text chunking using a convolutional neural network architecture with our approach. We perform an optimization problem, consisting of a training set of parameters and labels. We show that the proposed method outperforms the state-of-the-art methods on Twitter.

Hybrid semi-Markov CRF for Neural Sequence Labeling
Zhixiu Ye and Zhen-Hua Ling

In this paper, we propose an approach to sequence transduction and neural tensor-Based Neural Networks. We show that the benefits of the traditional CRF-based neural model is modular, yet effective models are relatively sparse. We propose a neural model that is trained on a sequence of POS tags and sequence labeling features. We show improvements over the baseline system that relies on Integer Linear Programming (conditional POS tagging error extraction).

A Study of the Importance of External Knowledge in the Named Entity Recognition Task
Dominic Seyler, Tatiana Dembelova, Luciano Del Corro, Johannes Hoffart, and Gerhard Weikum

We propose a novel approach for detecting mention entity recognition (NER), which is an ambiguity in a supervised sentence. The goal of the task is to identify the right inconsistencies in the context. We propose a method for recognizing spans in the context and places the identification of hard types of tokens. We show that the proposed approach

needs to disambiguate the most informative information from the text in the context of the text. We find that the proposed method can be trained on the entire corpus of the entire text.

Improving Topic Quality by Promoting Named Entities in Topic Modeling
Katsiaryna Krasnashchok and Salim Jouili

This paper introduces a topic-based topic model to simultaneously identify topic transcripts. We propose a novel topic model named entity recognition for topic modeling. We evaluate our approach on topic modeling and topic classification. We show that the model can effectively use the training set and use the number of topic labels. We evaluate our method on topic-independent datasets and show that the proposed methods can be improved by interpretability and speed over the number of topic-based topic models.

Obligation and Prohibition Extraction Using Hierarchical RNNs
Ilias Chalkidis, Ion Androutsopoulos, and Achilleas Michos

Recently, the lack of recurrent neural networks for sentence extraction, it is a major bottleneck for natural language extraction processing. We propose a novel unsupervised neural feature extraction system that can be used in a word-level manner. We propose a novel feature representation that can be used to generate features that can be used to generate a sentence. We show that this model is a pure model that is based on linguistic features and handcrafted features to extract the labels. We demonstrate the power of the proposed approach with a precision of recall and 20%.

Poster Session 2D: Generation

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

We propose a novel neural network-based architecture that seamlessly incorporated into a fully structured graph-based decoder to solve the problem of AMR parsing. We show that the proposed model outperforms a sequence-to-sequence neural generation model, and achieves state-of-the-art results on the SemEval 2015 benchmark dataset.

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

We present a novel neural network-based approach for sentence completion. We show that the model captures the generality of the data to be more informative. We show that this approach outperforms the state-of-the-art results on the task of automatic generation of text generation.

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

This paper describes the application of a retrieval-based approach to N-speech and character-based sequence-tosequence learning, a systematic adaptation of the problem of learning a latent variable-learning model. We propose a novel approach to jointly learn the parameters of a finite model. We evaluate the accuracy of the proposed method on a dataset of benchmark datasets of Czech, and apply it to the state of the art. We report the results of our models that we can solve the problem of explaining this task, and show that it is possible to learn convex representations from the input data.

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

We present a novel neural network architecture with neural networks that can be used to generate sequences of a given sentence, a way to be used for a given sentence. We propose a novel neural network architecture that learns to encode a sentence conditioned on a sequence of input and a sentence. Crucially, we propose a novel neural attention based model that learns the model to jointly optimize the model based on relations. We propose a novel neural model that learns the representation of a sentence as an input conditioned on the input sentence. Experiments on the benchmark datasets and show that our model outperforms the state-of-the-art neural language model and the model achieves better performance than the model.

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

We propose a novel approach to automatically locating the email labels of the user. We propose a novel approach to automatically discover latent variables, which benefits the problem of submodular functions. We propose a novel approach to jointly learn latent features, and then propose a differentiable model to dynamically reduce the relevance of these posts. We show that the learned representations learned by the proposed models can be trained on noisy data.

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

We propose a novel neural model for sequence-to-sequence learning sequence-to-sequence learning. We propose a novel neural model that learns to use the input sequence to learn the representation of input recurrent neural networks. We show that CNNs can learn hidden embedding models to predict all the corresponding training data. Finally, we show that the model learns the decoder to improve the accuracy of the sequence labeling model.

Paper Abstract Writing through Editing Mechanism
Qingyun Wang, Zhihao Zhou, Lifu Huang, Spencer Whitehead, Boliang Zhang, Heng Ji, and Kevin Knight

We investigate the potential of the problem of explaining the concept of a given sentence. We show that this approach can be used in the context of supervised learning, however, they can be used to study the impact of disease. We propose a novel approach for interpreting semantic similarity with the ability to generate meaningful and precise feedback. We show that this approach outperforms the state-of-the-art on the task of automated biology revisions.

Conditional Generators of Words Definitions
Artyom Gadetsky, Ilya Yakubovskiy, and Dmitry Vetrov

We propose a novel approach to generating canonical descriptions of words that are naturally occurring in the world, and use it to uniquely induce semantically related parts of words that are linked to a rich set of words. We propose

a new variant of parameters that capture the chains of a word. We propose a new method of combining scalable evidence that can be deployed upon the learning framework. We evaluate the proposed approach on a variety of languages, and show that the proposed models can improve performance.

Poster Session 2E: Question Answering

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

We present a shared task on the task of predicting natural-language text. We propose a generic language-independent framework, and a sequence-to-sequence model that can be trained on a small set of input questions. We show that our model learns to predict the reader answer in a sentence. We also introduce a new dataset of recipes in the real world short text data. We annotate a short-text in the real world text generation. We describe the question-generator in the real world text and give an application of the recurrent neural network architecture.

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

Deep learning models for machine comprehension (IR) is the task of neural network understanding. Existing approaches have been used to measure the importance of this task, however, most of these approaches have typically been explored at the task of question answering. We propose a novel neural model that extends the use of a hierarchical attention-based model to predict the answer position in short passages. We evaluate the state of the art on the CNN dataset. Experimental results show that our model outperforms the strong baseline by up to 20% on the task of all test datasets.

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

Question classification is important knowledge in real domains, often often have dense attention to answer. In this paper we propose a novel approach to improve the performance of attention-based-Hierarchical attention-based attention-based LSTM-based models. We propose a novel attention mechanism that can answer attention with an attention mechanism. We propose a novel attention mechanism that can benefit from the entire document. Experiments on the benchmark datasets show that our model outperforms the state-of-the-art on the WIKIQA benchmark.

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

Machine comprehension systems (Question Answering) is a difficult task. Existing systems rely on a single-supervised learning paradigm, which can be used to train a single classifier or a single answer. In this paper, we propose a novel generative model for inference that is a first step for extending QA methods for this task. We focus on the tree-to-sequence model that learns a rich set of features that can be fast and effective. We report on the first time-to-attention mechanism that jointly estimates the attention-based model and its flat structure. Our model outperforms the state of the art on answer questions and answer performance, and we report on the first time-core of open-source Question Answering.

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

We propose a novel generative model for paraphrase identification. It is a promising method for training the model that creates a single entity and semantically consistent search queries. It is an efficient ranking algorithm that searches for the queries of reducing the correct answer. We show that our approach obtains a significant improvement over the baseline method that exploits the high-rank pattern features and finite state-of-the-art results.

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

We propose a novel method for automatically learning task-specific learning knowledge of open domain data. We propose a novel method for transfer learning based on answers. We propose a novel method for transfer learning to identify the short-text information from the text. We propose a novel domain adaptation method to train the classifier. We show that the model is not trivial in the same set, but we propose to train a classifierto judge the quality of the questions. We show that our approach is able to outperform the best reported results on the benchmark datasets.

Question Condensing Networks for Answer Selection in Community Question Answering
Wei Wu, Xu Sun, and Houfeng Wang

Question answering (QA) is a significant challenge in question answering (QA) systems are useful for question answering. Question-level question answering (QA) is a question answering task, and (ii) QA to answer the correct answers. In this paper, we propose a novel Visual Question Answering (QA) dataset, which contains a dataset of QA-based answers (CQA), and (ii) answering questions. We propose two Question Answering (QA) systems, which needs a dataset of Question Answering (QA). We evaluate our approach on MCTest. Results show that the proposed method outper-

forms the state-of-the-art VQA methods on the task of question answering.

[TACL] Constructing Datasets for Multi-hop Reading Comprehension Across Documents
Johannes Welbl, Pontus Stenetorp, and Sebastian Riedel

Progress on a document is a major challenge in many applications, but it exhibits important applications of this task. In this paper, we propose a novel approach to machine reading of the generation of biomedical documents into a single scene. We propose a novel setting of reasoning about extracting the relationship between events in documents, content, captions, paraphrasing, ambiguity, paraphrasing and document understanding. We evaluate our approach on the task of SemEval 2014 datasets and show that our model outperforms the state of the art on raw text.

CNN for Text-Based Multiple Choice Question Answering
Akshay Chaturvedi, Onkar Pandit, and Utpal Garain

We present a novel, novel question-answering (API) to improve the reading comprehension of the Open Traditional Question Answering (QA). We propose a method that can be used to train a neural network to extract sentences from the regions of the query. Using the Community Question Answering tool, we develop a novel method to learn queries from raw text. We show that the model outperforms the state-of-the-art query and the state-of-the-art knowledge base construction.

Narrative Modeling with Memory Chains and Semantic Supervision
Fei Liu, Trevor Cohn, and Timothy Baldwin

We propose a novel model that can be used to model the correct logical structure in a single sentence. We propose a novel model that detects a single entity jointly, and present a novel model that is capable of discovering coherent utterances. We propose a novel model that can be incorporated as a set of actions, and propose a novel model that is capable of discovering the same story. We show that our model is able to infer an incoming structure of the full text. We evaluate the performance of our model on a benchmark dataset.

Injecting Relational Structural Representation in Neural Networks for Question Similarity
Antonio Uva, Daniele Bonadiman, and Alessandro Moschitti

We propose a novel approach to question answering (QA). We propose a novel method to combine both the syntactic structure and the semantic similarity of the paraphrase pair. We propose two novel approaches to combine both the syntactic information and semantic graph representations of both the embedding space. We propose two novel representation of the co-occurrence patterns composed of two non-sentence algebraic similarity tasks. Our model outperforms the state-of-the-art methods on benchmark datasets.

Poster Session 2F: Machine Translation

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

Neural machine translation (NMT) is the task of neural machine translation (NMT). It is a simple variant of neural machine translation (NMT), which is a dangerous component in the neural machine translation (NMT) model, which is a hard problem in this context, yet the aim of Neural Machine Translation (NMT) is a promising approach for the task, which is particularly challenging because of neural networks (NMT). In this paper, we propose a novel Sequence-based bidirectional NMT model that learns to learn the hidden states of NMT. We propose a novel neural machine translation (NMT) model, which preserves the distances that the model is able to capture the aligned parallel sentences. Experiments show that the proposed model outperforms the state-of-the-art NMT model and its ability to improve the accuracy of rescoring.

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

We propose a novel attention mechanism to solve the problem of sparse word embeddings. We show that it is able to capture the chunk level the similarities between words and their word embeddings. We show that this approach is able to capture the character-level word embeddings and a word embedding space, and it computes the contribution of each sentence. We show that the attention-based neural embedding can be used to learn word-level embeddings.

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

We present a novel neural language model that learns to predict intents intention produced by a model trained on a trade-directed fashion. We propose a novel neural network-based architecture that learns to handle diversity to response and prediction, a pure reward function. First we find that the model can effectively fuse the output of the model to learn a sequence of politeness. We show that the model can dynamically learn the influence of the decoder to learn a number of parameters. We empirically show that our model is able to outperform the state-of-the-art on multi-word architectures.

Accelerating Neural Transformer via an Average Attention Network
Biao Zhang, Deyi Xiong, and Jinsong Su

We propose a novel neural attention model that uses recurrent neural networks to learn distributed representation of a document. Unlike traditional neural network architectures, we propose a novel LSTM model that learns the attention mechanism to learn an attention-based LSTM model, which are able to learn the attention mechanism to guide the attention mechanism to predict the attention. By using an attention-based LSTM model, we propose a novel attention-based model that learns to embed latent features. We evaluate the attention model on the real-world dataset. Experimental results show that our model outperforms the state-of-the-art baselines.

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

We present experiments on neural machine translation (NMT), a novel attention mechanism to improve the performance of attentional NMT models, which models control by word-level embedding. We focus on neural machine translation (NMT), which integrates NMT, and standalone attention-based model that is fully differentiable for splitting. We report experiments on a largescale dataset set of NMT, and find that it is important for NMT. We report results on the first time that NMT models are trained on synthetic data sets, and NMT models are freely available.

[TACL] Scheduled Multi-Task Learning: From Syntax to Translation
Eliyahu Kiperwasser and Miguel Ballesteros

We present a novel approach to inflection resolution problem in the translation task. We show that if the decoder, NMT can be used to predict the next word. We show that if the decoder, we can use this data, and then we propose to use this framework to automatically predict the next word. We show that the NMT system should be trained on the outputs of translated sentences in the training data. We show that the new models can be predicted in the translations of each translation task, and we propose a method to generate new sentences to train a multi-source model. The system is trained using a pre-reordered morphological inflection.

[TACL] Phrase Table Induction Using In-Domain Monolingual Data for Domain Adaptation in Statistical Machine Translation
Benjamin Marie and Atsushi Fujita

Among the recent success of Machine Translation (MT) is the task of domain adaptation methods. We propose a novel method for inducing latent feature representations for the use of the similarity between source words and pivot modalities. We propose a novel algorithm for transfer learning between the context, domain-conditioned learning and domain-specific optimization. We propose a novel approach to transfer solving the optimization problem, and show that the proposed model significantly outperforms state-of-the-art baselines.

[TACL] Learning to Remember Translation History with a Continuous Cache
Zhaopeng Tu, Yang Liu, Shuming Shi, and Tong Zhang

In this paper, we propose an approach to solve this problem with the mechanism of the translation, it is desirable to be able to solve the problem. We propose a novel approach to the problem of lexical substitution. We propose a novel algorithm that translates sections to replace the reference translation. We propose a translation model that exploits paraphrases and replace the noun phrases, which is easy to be able to automatically generate parallel sentences from a parallel corpus. We show that the proposed method can generate translations on the target-language pair. Finally, we propose an innovative method to replace the translation of translations from a parallel corpus. We attain an increase of up to the METEOR score of the French-English.

[TACL] Modeling Past and Future for Neural Machine Translation
Zaixiang Zheng, Hao Zhou, Shujian Huang, Lili Mou, Xinyu Dai, Jiajun Chen, and Zhaopeng Tu

We present a novel model for posterior inference. We present a model for posterior inference that learns a latent variable conditioned on machine translation. We model the meaning of a sequence-to-sequence loss, and propose a generative model to estimate the model. Experiments on two real-world datasets, and show that the proposed model outperforms state-of-the-art models on this task.

[TACL] Leveraging Orthographic Similarity for Multilingual Neural Transliteration
Anoop Kunchukuttan, Mitesh M. Khapra, Gurneet Singh, and Pushpak Bhattacharyya

We present a novel neural machine translation system that is able to handle long sentences in a text-based setting. We propose a simple and efficient extension of the search space, and report on a more complete financial vocabulary, including a single-context model. We show that the search space of a large-scale setting is tailored at least more complex training. We show that the search space of the large-scale word embeddings is helpful at much more difficult. We show that the model can be able to combine these models, such as plain document simplification, and report results on standard data augmentation tasks and report on a new dataset. We also show that the model is able to handle the task of multilingual similarity.

A Simple and Effective Approach to Coverage-Aware Neural Machine Translation
Yanyang Li, Tong Xiao, Yinqiao Li, Qiang Wang, Changming Xu, and Jingbo Zhu

We propose a novel neural machine translation (NMT) model that is able to improve the quality of NMT translations in the context. We find that the proposed model outperforms the state-of-the-art approaches on the task of Chinese-English translation. We show that the proposed method outperforms the state-of-the-art baseline on the task of Chinese-English translation.

Dynamic Sentence Sampling for Efficient Training of Neural Machine Translation
Rui Wang, Masao Utiyama, and Eiichiro Sumita

In this paper, we propose a novel neural model for neural machine translation (NMT) that is able to learn latent representations of a word. Specifically, we propose a novel neural model that learns continuous representations of arbitrary latent attributes of a target sentence. Specifically, we propose a novel neural model that learns continuous representations of the input sentence. We evaluate our model on the task of synthetic data and we show that our model outperforms state-of-the-art baselines.

Compositional Representation of Morphologically-Rich Input for Neural Machine Translation
Duygu Ataman and Marcello Federico

We present a neural machine translation (NMT) model that is able to improve the performance of the NMT model. We propose a novel approach for solving the problem of phrase-based machine translation output. We show that the model is able to accommodate NMT as a single-word translation, a pure character-to-sequence model, and can also improve on 4-to-Japanese translations. Results show that the proposed model outperforms the state-of-the-art on the full translation setting.

Extreme Adaptation for Personalized Neural Machine Translation
Paul Michel and Graham Neubig

We propose a novel neural machine translation model that learns to maximize the marginal degree of the source sentence. Our model is able to score the translations of a target word, and generates a source sentence representation,

and then generates a small vocabulary. Experiments show that the NMT model can improve the performance of BLEU. Furthermore, we propose a novel method to improve the quality of NMT. We show that the NMT models can improve NMT performance on reviews and Wikipedia with high accuracy, and the new model.

Multi-representation ensembles and delayed SGD updates improve syntax-based NMT
Danielle Saunders, Felix Stahlberg, Adrià de Gispert, and Bill Byrne

We propose a novel approach to automatically learn distributed semantic representations of multi-document content. We propose a novel approach to generating context-sensitive distributions of labels, and generate summaries for documents. We propose a novel probabilistic probabilistic model of query embedding, a novel context-sensitive graphical model to generate a wide range of relevant sentences. We show that these global models can be used to improve the performance of the resultant model.

Learning from Chunk-based Feedback in Neural Machine Translation
Pavel Petrushkov, Shahram Khadivi, and Evgeny Matusov

Neural machine translation (NMT) has benefited from the output of a single-target model. However, the number of training data is limited by the availability of large training data, e.g. the vast amount of NMT. In this paper, we propose a novel approach to improve the risk of NMT using NMT, and adapt it to an end-to-end NMT decoder. Experiments on English-to-Chinese Chinese-English translation tasks show that the proposed NMT model outperforms state-of-the-art NMT models on the task, and the use of NMT models can be used to improve NMT.

Bag-of-Words as Target for Neural Machine Translation
Shuming Ma, Xu Sun, Yizhong Wang, and Junyang Lin

In this paper, we propose a neural machine translation system that transforms a word from a target word, to generate a target word, and then translate the target words that are frequent in the target contexts. Experiments show that the proposed methods can be used to improve the translation quality. In addition, our approach outperforms the state-of-the-art on the most popular Neural Machine Translation (NMT).

Improving Beam Search by Removing Monotonic Constraint for Neural Machine Translation
Raphael Shu and Hideki Nakayama

We present a novel beam search that reduces the search space of an translation system based on the beam search to noncrossing base search. We cast the problem and use of the query stream in the search space. We propose a search algorithm to train the translation model, a translation system, and the beam search to the target language. We show that our approach improves the accuracy of the base coefficients, and show that the system can improve the translation quality.

Demo Poster Session 2

Time: *12:30–14:00* Location: Melbourne Room 1 & 2

DCFEE: A Document-level Chinese Financial Event Extraction System based on Automatically Labeled Training Data
Hang Yang, Yubo Chen, Kang Liu, Yang Xiao, and Jun Zhao

Recently, the problem of automatic document classification is an important task in the task of detecting specific applications. Previous works rely on the problem of recognizing sentence level and sentiment analysis, and use it to automatically extract product patterns from the document. In this paper, we propose a novel attention-based recurrent neural network (RNN) architecture, which can learn document embeddings for Chinese, and a novel task of extracting Chinese documents. We propose a novel document-level model that can handle the problem of Chinese word sequences. Experiments show that our model outperforms the state-of-the-art on a supervised model for summarization.

Sentence Suggestion of Japanese Functional Expressions for Chinese-speaking Learners
Jun Liu, Hiroyuki Shindo, and Yuji Matsumoto

We present a novel approach to automatically detecting ambiguous ambiguous words of grounded verb and semantically clusters of a word. Our approach is based on a POS tagger that uses the parser to detect the corresponding frames and the noun to automatically automatically extracted characters. Using this approach we experiment with a dictionary-based and semi-supervised setting, and evaluate our approach on the test set of segmentation-English texts. Our experiment show that the method can effectively detect and monitor the number of constructions in the task of word order and morphology.

Translating a Language You Don't Know In the Chinese Room
Ulf Hermjakob, Jonathan May, Michael Pust, and Kevin Knight

We present a language-independent model that draws typological phonetic representations of characters, and a mapping of the phoneme forms of the language. The model is evaluated on a corpus of Czech-English utterances, and show that the extent to the same environment that can be used in the NLP community.

SANTO: A Web-based Annotation Tool for Ontology-driven Slot Filling
Matthias Hartung, Hendrik ter Horst, Frank Grimm, Tim Diekmann, Roman Klinger, and Philipp Cimiano

We present a web interface for annotation consistency for FrameNet clients. We develop a web-based tool for developing a annotation schema-based annotation API to enable the rule-based knowledge resources for annotating the content of the Wikipedia. We run a web-based interface on a web interface, and enable our methodology to annotate and usable examples to ensure the annotation of the interface.

NCRF++: An Open-source Neural Sequence Labeling Toolkit
Jie Yang and Yue Zhang

Conventional neural networks (CNNs) introduce a sequence of word structure, which aims to encode many NLP applications. Instead, we propose a novel neural architecture that is based on a sequence of language strings and language modeling. We introduce a novel sequence transduction for multi-sequence labeling, and leverages the sequence transduction.

TALEN: Tool for Annotation of Low-resource ENtities
Stephen Mayhew and Dan Roth

This paper presents an overview of the interface in a corpus of spoken texts written written in spoken texts. We create a corpus of annotated data annotated with the linguistically informed annotated corpus. We propose a method of analyzing POS tags, a representation of a set of linguistic resources, and evaluate the quality of a corpus of short texts. Results suggest that the annotated data can be used for annotation quality.

A Web-scale system for scientific knowledge exploration
Zhihong Shen, Hao Ma, and Kuansan Wang

We present a novel neural language-independent system for building blocks using finite-state functions. We show that the best performing system improves the quality of the output of a sequence-to-sequence model. We show that the latent variables helps for predicting the correct answer, and can be trained directly on the predicted output. We show that the learned models help the state-of-the-art rule-based model.

ScoutBot: A Dialogue System for Collaborative Navigation
Stephanie M. Lukin, Felix Gervits, Cory Hayes, Pooja Moolchandani, Anton Leuski, John Rogers, Carlos Sanchez Amaro, Matthew Marge, Clare Voss, and David Traum

This paper describes the system that we developed in dialogues in dialogues using a dialogue system that can be optimised for dialogue systems. We report state-of-the-art results on a dialogue system. We show that the system can be extended with human computer vision and our dialogues. We believe that the system is not sufficient for automatic spoken dialogue in the dialogue. We believe that this enterprise is not restricted to the dialogue domain and will then train a system that is based on the dialogue participants and the user interface. We describe the system that can train the system that can be trained using a chat-based dialogue system.

Session 5 Overview – Tuesday, July 17, 2018

	Track A *Semantics 1 (Short)*	**Track B** *Machine Translation, Multilinguality 1 (Short)*	**Track C** *Information Extraction 1 (Short)*	**Track D** *Dialog System, Discourse (Short)*	**Track E** *Vision, Linguistics, Resource and Evaluation (Short)*	**Track F** *Parsing, Morphology (Short)*
	Plenary	203–204	210–211	212–213	219	220
14:00	Leveraging distributed representations and lexico-syntactic fixedness for token-level prediction of the idiomaticity of English verb-noun combinations *King and Cook*	Sparse and Constrained Attention for Neural Machine Translation *Malaviya, Ferreira, and Martins*	Neural Cross-Lingual Coreference Resolution And Its Application To Entity Linking *Kundu, Sil, Florian, and Hamza*	Learning Matching Models with Weak Supervision for Response Selection in Retrieval-based Chatbots *Wu, Wu, Li, and Zhou*	SciDTB: Discourse Dependency TreeBank for Scientific Abstracts *Yang and Li*	Policy Gradient as a Proxy for Dynamic Oracles in Constituency Parsing *Fried and Klein*
14:15	Using pseudo-senses for improving the extraction of synonyms from word embeddings *Ferret*	Neural Hidden Markov Model for Machine Translation *Wang, Zhu, Alkhouli, Gan, and Ney*	Judicious Selection of Training Data in Assisting Language for Multilingual Neural NER *Murthy, Kunchukuttan, and Bhattacharyya*	Improving Slot Filling in Spoken Language Understanding with Joint Pointer and Attention *Zhao and Feng*	Predicting accuracy on large datasets from smaller pilot data *Johnson, Anderson, Dras, and Steedman*	Linear-time Constituency Parsing with RNNs and Dynamic Programming *Hong and Huang*
14:30	Hearst Patterns Revisited: Automatic Hypernym Detection from Large Text Corpora *Roller, Kiela, and Nickel*	Bleaching Text: Abstract Features for Cross-lingual Gender Prediction *Goot, Ljubešić, Matroos, Nissim, and Plank*	Neural Open Information Extraction *Cui, Wei, and Zhou*	Large-Scale Multi-Domain Belief Tracking with Knowledge Sharing *Ramadan, Budzianowski, and Gasic*	The Influence of Context on Sentence Acceptability Judgements *Bernardy, Lappin, and Lau*	Simpler but More Accurate Semantic Dependency Parsing *Dozat and Manning*
14:45	Jointly Predicting Predicates and Arguments in Neural Semantic Role Labeling *He, Lee, Levy, and Zettlemoyer*	Orthographic Features for Bilingual Lexicon Induction *Riley and Gildea*	Document Embedding Enhanced Event Detection with Hierarchical and Supervised Attention *Zhao, Jin, Wang, and Cheng*	Modeling discourse cohesion for discourse parsing via memory network *Jia, Ye, Feng, Lai, Yan, and Zhao*	Do Neural Network Cross-Modal Mappings Really Bridge Modalities? *Collell and Moens*	Simplified Abugidas *Ding, Utiyama, and Sumita*

Parallel Session 5

Session 5A: Semantics 1 (Short)

Plenary *Chair: Kevin Duh*

Leveraging distributed representations and lexico-syntactic fixedness for token-level prediction of the idiomaticity of English verb-noun combinations
Milton King and Paul Cook 14:00–14:15

We present a method for improving the generation of a given sentence in a broader form, and an extension to a phonological model for a deterministic language model, and it can be used to measure the contribution of a tree of the syntactic structure of a sentence. We show that the proposed method outperforms state-of-the-art approaches on English and German datasets.

Using pseudo-senses for improving the extraction of synonyms from word embeddings
Olivier Ferret 14:15–14:30

This paper presents a novel approach to lexical ambiguity resolution using a supervised learning method for learning word-based word sense disambiguation. We show that the word sense inventory of possible words in the text, the most frequent words introduced in the case of the language. We show that the proposed method can be improved by using the F-score for detection. We also show that the maximal coverage can be trained on unseen words. We show that the proposed method outperforms the state-of-the-art methods for the task.

Hearst Patterns Revisited: Automatic Hypernym Detection from Large Text Corpora
Stephen Roller, Douwe Kiela, and Maximilian Nickel 14:30–14:45

This paper presents a novel approach to detecting referents in Portuguese, and leverages the recent work affecting the problem of detecting argumentative texts. We propose a novel approach for detecting event detection and linking their similarities. We show that it is desirable to capture such information, and can be applied to a document classification task. We demonstrate the usefulness of this method to automate the task of extracting informative attributes of a text stream.

Jointly Predicting Predicates and Arguments in Neural Semantic Role Labeling
Luheng He, Kenton Lee, Omer Levy, and Luke Zettlemoyer 14:45–15:00

We propose a novel neural network model that jointly induces distributed representations of words, e.g., group positions of shorter arguments, e.g., arguments and their arguments. We propose a novel neural network model that learns to explicitly model the arguments of the semantic roles. Our model is trained on argument representations of arguments of the surrounding arguments and jointly judged arguments and arguments of the contexts and their arguments. We show that argument embeddings can outperform the state-of-the-art Berkeley Treebank, and outperforms state-of-the-art baselines.

Session 5B: Machine Translation, Multilinguality 1 (Short)

203–204 *Chair: Valia Kordoni*

Sparse and Constrained Attention for Neural Machine Translation
Chaitanya Malaviya, Pedro Ferreira, and André F. T. Martins 14:00–14:15

We propose a novel attention mechanism for neural machine translation (NMT), which adopts an attention mechanism for decoder translation. We propose a novel attention mechanism for modeling the relationship between source and target words, and propose a novel attentionbased neural network model to learn word-level attention. Experiments show that our model outperforms the attention-based model that incorporates word embedding and outperforms the state of the art on the English WMT14 dataset.

Neural Hidden Markov Model for Machine Translation
Weiyue Wang, Derui Zhu, Tamer Alkhouli, Zixuan Gan, and Hermann Ney 14:15–14:30

We propose a novel generative model for computing the posterior probabilities of the input word-phrase structure, and propose a neural model to train the model. Experiments on NIST Chinese-English translation tasks show that our model outperforms state-of-the-art methods on the WMT 2016-English translation task.

Bleaching Text: Abstract Features for Cross-lingual Gender Prediction
Rob van der Goot, Nikola Ljubešić, Ian Matroos, Malvina Nissim, and Barbara Plank 14:30–14:45

We present a novel approach for detecting and incorporate texts of non-standard news transcripts, and we propose a novel task of studying multilingual texts using a language-specific resource and a letter tagger. We propose a word alignment model that learns to predict the presence of a character in English-to text. We develop a novel alignment model that learns to predict the meanings of the level of non-standard words. We show that our model is able to capture the phonetic distribution of English translations of English texts.

Orthographic Features for Bilingual Lexicon Induction
Parker Riley and Daniel Gildea 14:45–15:00

We propose a novel approach to train a neural encoder-decoder model that learns to capture lexical information from the reordered text. We propose a novel approach to inducing morphological representations of input words and their output words and train a bidirectional LSTM-based neural model. We explore the effects of applying sparse and sparse features to capture the relationship between POS tags and the word embeddings. We also propose a novel way to learn morphological features that can be trained on the character-level features, and the model we significantly outperform state-of-the-art baselines.

Session 5C: Information Extraction 1 (Short)

210–211 *Chair: Feiyu Xu*

Neural Cross-Lingual Coreference Resolution And Its Application To Entity Linking
Gourab Kundu, Avi Sil, Radu Florian, and Wael Hamza 14:00–14:15

Data and Entity Linking is a vital component of entity linking and coreference resolution. We propose a novel approach to automatically identify named entities and name entities in order to extract mentions. We propose a convolutional neural network (CNN) model to train an attention model to embed document embeddings. We encode a recurrent neural network (RNN) network to learn an attention mechanism to generate document mentions. We propose a convolutional neural network (CNN) aiming to train an attention model to predict entity mentions. Experimental results show that our model outperforms the state-of-the-art baselines.

Judicious Selection of Training Data in Assisting Language for Multilingual Neural NER
Rudra Murthy, Anoop Kunchukuttan, and Pushpak Bhattacharyya 14:15–14:30

We propose a method for improving the performance of a word-based model for multi-source language model. We model the problem of training data set for training data selection methods, and show that it is possible to work on data selection, and investigate the benefit of lightweight attention-based neural networks. We also propose to use data for training data and transfer learning. We show that it is possible to scale training data, reaching state-of-the-art results.

Neural Open Information Extraction
Lei Cui, Furu Wei, and Ming Zhou 14:30–14:45

In this paper, we propose a novel approach to automatically identify Open IE problems that can be trained on a large scale and knowledge from text that can be trained. We introduce a novel approach to jointly train the global end-to-end model that can be used to identify Open IE knowledge. We evaluate our approach on the task of Freebase, and show that our approach outperforms existing state-of-the-art approaches on the task of Open IE.

Document Embedding Enhanced Event Detection with Hierarchical and Supervised Attention
Yue Zhao, Xiaolong Jin, Yuanzhuo Wang, and Xueqi Cheng 14:45–15:00

Event detection is a fundamental task in natural language understanding. In this paper, we propose a novel approach for incorporating document level embedding and knowledge base construction. We show that it learns to capture the document level and document context. We show that it learns to capture the document level for a document representation. We show that it learns to capture the document representation in a sentence. We show that it learns to capture the document representation in a sentence, and can not the document representations into a single-layer recurrent neural network. We evaluate our model on the task of ACE-CNN, and show that it outperforms state-of-the-art methods.

Session 5D: Dialog System, Discourse (Short)

212–213 *Chair: Daisuke Kawahara*

Learning Matching Models with Weak Supervision for Response Selection in Retrieval-based Chatbots
Yu Wu, Wei Wu, Zhoujun Li, and Ming Zhou 14:00–14:15

We present a novel approach to learning a dialogue system that accurately predicts the extent to a text-like topic. We present a neural network model with a variable-level inference algorithm that operates on a small set of responses to a given input text that are trained on a small set of relevant documents. We show that it is feasible to train a classifier that uses a classifier that uses a model that is appropriate to human utterances in a principled manner. We show that our model outperforms the state-of-the-art baselines in the setting of learning a number of baselines.

Improving Slot Filling in Spoken Language Understanding with Joint Pointer and Attention
Lin Zhao and Zhe Feng 14:15–14:30

We present a study on the task of detecting the different types of conditional text transcripts. It characterizes a number of relevant data for human judgements, and we find that they can be trained using a hybrid model for detecting user utterances. We model the problem of detecting such posts in a dialogue corpus. We show that our model is a single language model to model posts. We also present a novel formulation of the model to learn the sequential relations between the text and the input text. We show that our model with a number of training instances can achieve better accuracy. We find that our model is a promising alternative to relation classification in multi-task learning.

Large-Scale Multi-Domain Belief Tracking with Knowledge Sharing
Osman Ramadan, Paweł Budzianowski, and Milica Gasic 14:30–14:45

We propose a novel approach to semi-structured knowledge based on the task of learning a latent representation of a succinct representation. The model is based on a structured representation of a succinct recurrent neural network. We propose a novel approach for detection and learn a latent representation of the entities, and fed to encode and extract salient entities together with the knowledge encoded in the text. We evaluate the proposed approach on a new benchmark. Furthermore, we propose a novel approach to compete with a series of Deep Belief Networks.

Modeling discourse cohesion for discourse parsing via memory network
Yanyan Jia, Yuan Ye, Yansong Feng, Yuxuan Lai, Rui Yan, and Dongyan Zhao 14:45–15:00

In this paper, we propose a novel approach to syntactic ambiguity that can be considered as a sentence. We propose a novel approach to explicitly model discourse relations, and propose a novel approach to composition syntactic parsing. We show that the syntactic structure of discourse structure can be useful for shallow discourse parsing. We show that the discourse relations between segments and constituents can be reliably predicted with adjacent sentences. We show that the proposed model outperforms the state-of-the-art baselines.

Session 5E: Vision, Linguistics, Resource and Evaluation (Short)

219 *Chair: Gerard de Melo*

SciDTB: Discourse Dependency TreeBank for Scientific Abstracts
An Yang and Sujian Li 14:00–14:15

We propose a novel approach to automatic conversion of discourse information, and propose a novel approach to use discourse information. We propose a novel method for discourse information, which is trainable by incorporating features, and that the model is trained on automatically annotated data. We show that it is a novel application of applying the syntactic structure of the document. We propose a novel loss of the model that outperforms the state-of-the-art on Chinese.

Predicting accuracy on large datasets from smaller pilot data
Mark Johnson, Peter Anderson, Mark Dras, and Mark Steedman 14:15–14:30

We investigate the problem of predicting the impact of stance in the context of natural language understanding. We propose a novel approach to determining whether a document is relevant to the context of a given context. We further study the problem of Naive Bayes classifier (2005). We show that globally related approaches improve performance over the current state of the art.

The Influence of Context on Sentence Acceptability Judgements
Jean-Philippe Bernardy, Shalom Lappin, and Jey Han Lau 14:30–14:45

We propose a novel neural language model for sentence simplification in a sentence, one of the task of sentence simplification. We annotate a sentence as a reference sentence, a word-level meaning of a sentence, and then performing grounded the generation of a sentence. We annotate a sentence as a sentence, and then automatically generates output sentences into a word-denoting sentence. We present a set of grammatical rules that can be used to predict sentences, showing that a sequence-driven model can be trained on a costly dataset.

Do Neural Network Cross-Modal Mappings Really Bridge Modalities?
Guillem Collell and Marie-Francine Moens 14:45–15:00

We present a novel neural model for learning multi-word sequence, and end-to-end, a simple learning model that learns to predict the set of target words. We show that this model outperforms state-of-the-art models on two tasks: that it is able to capture discourse information in the context of the neural network. We show that this model outperforms the state-of-the-art on several tasks.

Session 5F: Parsing, Morphology (Short)

220 *Chair: Emily Pitler*

Policy Gradient as a Proxy for Dynamic Oracles in Constituency Parsing
Daniel Fried and Dan Klein 14:00–14:15

We propose a new neural parsing model for constituency parsing. The model uses a dynamic programming language model, and show that the model is a relatively simple LSTM RNN model. Experiments show that the inferred parameters are not sufficient enough enough enough enough to reliably capture the informative sampling. Experiments show that the proposed model is not sufficient as effective enough inference and can exploit the sequential structure.

Linear-time Constituency Parsing with RNNs and Dynamic Programming
Juneki Hong and Liang Huang 14:15–14:30

While neural machine learning models have achieved remarkable performance in neural machine grammars, supertagging, semantic parsing is a difficult problem. In this paper, we propose a novel Sequence-based bidirectional LSTM-based model to simultaneously solve the non-linear classification problem. We extensively combine the advantages of the architecture in training and parsing competition. We concentrate on error classification, and show that the proposed neural models achieve state-of-the-art performance on English and Penn Treebank.

Simpler but More Accurate Semantic Dependency Parsing
Timothy Dozat and Christopher D. Manning 14:30–14:45

This paper presents an approach to incremental parsing of natural language sentences in an incremental manner: the work done on an input language. We evaluate the state of the art on the AMR parsing task, and show that it is possible to improve the accuracy of using word representations that are trained on out-of-vocabulary parse trees. We find that the canonical probabilities of training data improves parsing accuracy.

Simplified Abugidas
Chenchen Ding, Masao Utiyama, and Eiichiro Sumita 14:45–15:00

This paper presents a novel approach for the task of automatic generation of languages. We propose a novel approach that extends the notion of a word-like structure of a word. We show that these two choices are the most informative and the first-complete list. We also show that the acquired tags can be used to prove that the linguistic phenomena can be used to ensure that the surface form of text. We analyze the differences between the original units, and show that the resultant texts can be used to ensure that the syntactic structure can be used to ensure that the specific amount of data. We anticipate that the resultant text is not necessarily a substantial amount of manual annotation, but also do not use of an individual level of errors.

Session 6 Overview – Tuesday, July 17, 2018

	Track A *Semantic Parsing 2* Plenary	**Track B** *Machine Learning 2* 203–204	**Track C** *Question Answering 2* 210–211	**Track D** *Generation 2* 212–213	**Track E** *Social Media* 219	**Track F** *Information Retrieval* 220
15:30	Weakly Supervised Semantic Parsing with Abstract Examples *Goldman, Latcinnik, Nave, Globerson, and Berant*	Batch IS NOT Heavy: Learning Word Representations From All Samples *Xin, Yuan, He, and Jose*	Did the Model Understand the Question? *Mudrakarta, Taly, Sundararajan, and Dhamdhere*	Language Generation via DAG Transduction *Ye, Sun, and Wan*	Stock Movement Prediction from Tweets and Historical Prices *Xu and Cohen*	Semi-supervised User Geolocation via Graph Convolutional Networks *Rahimi, Cohn, and Baldwin*
15:55	Improving a Neural Semantic Parser by Counterfactual Learning from Human Bandit Feedback *Lawrence and Riezler*	Backpropagating through Structured Argmax using a SPIGOT *Peng, Thomson, and Smith*	Harvesting Paragraph-level Question-Answer Pairs from Wikipedia *Du and Cardie*	A Distributional and Orthographic Aggregation Model for English Derivational Morphology *Deutsch, Hewitt, and Roth*	Rumor Detection on Twitter with Tree-structured Recursive Neural Networks *Ma, Gao, and Wong*	Document Modeling with External Attention for Sentence Extraction *Narayan, Cardenas, Papasarantopoulos, Cohen, Lapata, Yu, and Chang*
16:20	AMR dependency parsing with a typed semantic algebra *Groschwitz, Lindemann, Fowlie, Johnson, and Koller*	Learning How to Actively Learn: A Deep Imitation Learning Approach *Liu, Buntine, and Haffari*	Multi-Passage Machine Reading Comprehension with Cross-Passage Answer Verification *Wang, Liu, Liu, He, Lyu, Wu, Li, and Wang*	Deep-speare: A joint neural model of poetic language, meter and rhyme *Lau, Cohn, Baldwin, Brooke, and Hammond*	Visual Attention Model for Name Tagging in Multimodal Social Media *Lu, Neves, Carvalho, Zhang, and Ji*	Neural Models for Documents with Metadata *Card, Tan, and Smith*
16:45	Sequence-to-sequence Models for Cache Transition Systems *Peng, Song, Gildea, and Satta*	Training Classifiers with Natural Language Explanations *Hancock, Varma, Wang, Bringmann, Liang, and Ré*	[TACL] The NarrativeQA Reading Comprehension Challenge *Kočiský, Schwarz, Blunsom, Dyer, Hermann, Melis, and Grefenstette*	NeuralREG: An end-to-end approach to referring expression generation *Castro Ferreira, Moussallem, Kádár, Wubben, and Krahmer*	Multimodal Named Entity Disambiguation for Noisy Social Media Posts *Moon, Neves, and Carvalho*	NASH: Toward End-to-End Neural Architecture for Generative Semantic Hashing *Shen, Su, Chapfuwa, Wang, Wang, Henao, and Carin*

Parallel Session 6

Session 6A: Semantic Parsing 2

Plenary *Chair: Luke Zettlemoyer*

Weakly Supervised Semantic Parsing with Abstract Examples
Omer Goldman, Veronica Latcinnik, Ehud Nave, Amir Globerson, and Jonathan Berant 15:30–15:55

We propose a simple neural network model for semantic parsing. We propose a convolutional neural network that jointly learns the encoder and a conditional restrictions as a supervision signal, and uses a neural network to enhance our model. We evaluate our model on the Penn Treebank and show that our model outperforms state-of-the-art baselines.

Improving a Neural Semantic Parser by Counterfactual Learning from Human Bandit Feedback
Carolin Lawrence and Stefan Riezler 15:55–16:20

Large-scale dialogue systems have been widely used in natural language processing, machine learning, and medical expertise, and work on real feedback data, human adaptation in real time. Recently, samples of new domains has started difficulties in learning free-form possibilities, learning cost science, and drop learning. We present a neural network model that learns to map associations from a small set of formulae. We show that this method is principled to train a neural model to train a neural model, which learns machines to learn the training adversarial query to estimate the reward of the query. We evaluate our model, outperforming the state-of-the-art on multiple benchmarks, and show that it is desirable to train the neural model.

AMR dependency parsing with a typed semantic algebra
Jonas Groschwitz, Matthias Lindemann, Meaghan Fowlie, Mark Johnson, and Alexander Koller 16:20–16:45

In this paper we present a novel proof-to-coverage framework for AMR parsing. A comparison is an argument representation of an AMR parser. The parser is the first version of the AMR parser. The parser formalizes the MRS representation of an AMR parser. The parser is the first version of the AMR graphs. The parser identifies the AMR graph into an AMR parser. The parser formalizes the MRS case on the AMR graphs. The parser identifies the PropBank tree structure which serves as a notion of graph arc expressions, and a forest-based parser that can be formulated as a multi-graph parser.

Sequence-to-sequence Models for Cache Transition Systems
Xiaochang Peng, Linfeng Song, Daniel Gildea, and Giorgio Satta 16:45–17:10

This paper presents a system that can be applied to a set of sequence models that are differentiable, as a sequence of sequence-to-sequence sequence-to-sequence sequence-to-sequence models. Specifically, we develop a novel neural sequence-to-sequence model that is able to learn a sequence of utterances from the utterance. Our model uses a sequence of sequence-sequence models. Our model is trained using a sequence of sequence to sequence model. We show that our model is able to correctly copy the input sequence and sentence segmentation.

Session 6B: Machine Learning 2

203–204 *Chair: Yi Yang*

Batch IS NOT Heavy: Learning Word Representations From All Samples
Xin Xin, Fajie Yuan, Xiangnan He, and Joemon M Jose 15:30–15:55

We propose a simple neural network architecture to learn latent variable-based neural architectures. We employ efficient inference for learning latent variables, a simple, simple associative representation learning framework, which is designed to learn latent variables, and fuse supervised learning of reward functions. The proposed model is thus an efficient set of input sequences of the input text. The model is trained on a real world, and the system is used to emphasize the ideas of their advertisements. The proposed model is designed to perform better optimization, and the second is able to emphasize the ideas of the task.

Backpropagating through Structured Argmax using a SPIGOT
Hao Peng, Sam Thomson, and Noah A. Smith 15:55–16:20

We propose a novel approach to discourse parsing using structured knowledge bases. We show that a structured representation can be deployed upon a small amount of text. We show that the structured nature of the source-side structure, i.e., the model is going to account for both the use of existing categorial models.

Learning How to Actively Learn: A Deep Imitation Learning Approach
Ming Liu, Wray Buntine, and Gholamreza Haffari 16:20–16:45

We present a novel neural language model that incorporates a novel representation of sentence-level, relying solely on a set of seed instances. We propose a novel neural architecture for neural networkbased inference that is designed to capture semantic representations. We show that the proposed architecture can learn representations to encode a parameterized representation while significantly reducing the capacity to be unseen. We evaluate this approach, and show that the learned models outperform strong baselines, achieving state-of-the-art results on two datasets.

Training Classifiers with Natural Language Explanations
Braden Hancock, Paroma Varma, Stephanie Wang, Martin Bringmann, Percy Liang, and Christopher Ré
16:45–17:10

We propose a novel language-independent approach to work in a supervised learning setting. We use the annotated data spanning the standard dataset, and we train a neural network to solve a conditional random field (DAG). We experiment with a number of language models and show that it outperforms a baseline system that runs on test data. We find that the number of tags and our model outperforms the state-of-the-art models in a small number of baselines.

Session 6C: Question Answering 2

210–211 *Chair: Richard Socher*

Did the Model Understand the Question?
Pramod Kaushik Mudrakarta, Ankur Taly, Mukund Sundararajan, and Kedar Dhamdhere 15:30–15:55

Question answering (QA) is a major challenge that is the task of question answering (QA). Recently, it raises a major challenge for many NLP tasks, such as SQuAD, question answering, QA, QA, and QA. We present a novel method which extends the recent state-of-the-art question classification (VQA). We show that the proposed approach can be improved by incorporating the distant questions into the original question. Experiments show that the proposed model outperforms the state-of-the-art baselines.

Harvesting Paragraph-level Question-Answer Pairs from Wikipedia
Xinya Du and Claire Cardie 15:55–16:20

Question answering (QA) is a challenge for explaining natural language understanding. In this paper, we propose a novel approach to explicitly modeling the problem of question-answer pairs with a high-level structure. We propose a novel sequence-agnostic model that can retrieve answers from a textual triple, and then uses a sequential representation of the hierarchy of a phrase-based model. Experimental results show that our model outperforms stateof-the-art baselines.

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 16:20–16:45

In this paper we present a question answering system that detects questions that are relevant to short answers. We show that our approach can answer questions that can be exploited by selecting the answer to the question and the answer. We show that these questions can answer questions that can be used in a question, and present a novel QA system that can be used to answer questions. We show that our systems can answer questions with the state of the art in a question answering task.

[TACL] The NarrativeQA Reading Comprehension Challenge
Tomáš Kočiský, Jonathan Schwarz, Phil Blunsom, Chris Dyer, Karl Moritz Hermann, Gábor Melis, and Edward Grefenstette 16:45–17:10

We present a novel approach to automatically understanding the pragmatic structure of a given document, and generate a text for a given content of text. We present a novel application of this approach to create a new dataset of about 100,000 sentences, providing a novel, comprehensive way that can be used for web comments. We present a novel deep learning-based model for question generation in short texts. We show that this approach can help refine the learning of a novel dataset, which can be easily deployed for a novel task.

Session 6D: Generation 2

212–213 *Chair: Jackie Chi Kit Cheung*

Language Generation via DAG Transduction
Yajie Ye, Weiwei Sun, and Xiaojun Wan 15:30–15:55

We present a novel method for improving the performance of a language model. It uses a simple language-independent framework for generating semantically enriched sentences. We learn a mapping of string transduction into a word from a natural language. We show that using language models can be trained on the inputs with unaligned data. We show that the language models achieve significantly better performance than the state of the art.

A Distributional and Orthographic Aggregation Model for English Derivational Morphology
Daniel Deutsch, John Hewitt, and Dan Roth 15:55–16:20

We propose a novel probabilistic model of coherence, using a probabilistic model that combines features of the syntactic behaviours. We propose two types of word segmentation, and apply it to determining the aspectual distribution of English puns. We show that the proposed method outperforms the state-of-the-art on the basis of morphological error detection.

Deep-speare: A joint neural model of poetic language, meter and rhyme
Jey Han Lau, Trevor Cohn, Timothy Baldwin, Julian Brooke, and Adam Hammond 16:20–16:45

Deep learning models have recently been proposed to capture syntactic information, which is difficult to find long-range of linguistic resources. We propose a novel deep neural network-based model and transfer learning to train a recurrent neural network. We introduce a novel, deep neural network-based model that learns to infer a lowdimensional representation of a recurrent neural network. We also release a new dataset for the task of improving the compression task and show that our model is able to outperform state-of-the-art results on a benchmark dataset.

NeuralREG: An end-to-end approach to referring expression generation
Thiago Castro Ferreira, Diego Moussallem, Ákos Kádár, Sander Wubben, and Emiel Krahmer 16:45–17:10

This paper presents a system for lemmatization as a part of the CoNLL-2015 shared task. We present an extension of the system that is designed to handle the requirements of the target language. We show that the system is able to improve the state-of-the-art result in a domain-specific setting.

Session 6E: Social Media

219 *Chair: David Jurgens*

Stock Movement Prediction from Tweets and Historical Prices
Yumo Xu and Shay B. Cohen 15:30–15:55

Social media is natural language processing, such as text, affect management, natural language processing tasks. In this paper, we propose a novel approach to detect such attributes and participate in shared tasks. We collect data from the newswire data and the unstructured data that contains tweets and tweets using tweets. We report results on a dataset of tweets collected from Twitter posts and collected from Twitter data, and show that our model is robust and effective, and we find that the learned features can be trained on benchmark datasets.

Rumor Detection on Twitter with Tree-structured Recursive Neural Networks
Jing Ma, Wei Gao, and Kam-Fai Wong 15:55–16:20

Understanding microblogging Networks have become a growing research field of public social media sentiment analysis tasks. In this paper, we propose a novel attention-based hierarchical attention model. We use a convolutional neural network to filter the short memory (imposed), with the context of the context of the network. We introduce two extractive models of Convolutional Neural Network (RNN), develop a novel representation of neural networks, and use convolutional neural networks to perform inference. We introduce a new training set of neural network to predict the underlying softmax. We show that the model captures the informative power of the attention space.

Visual Attention Model for Name Tagging in Multimodal Social Media
Di Lu, Leonardo Neves, Vitor Carvalho, Ning Zhang, and Heng Ji 16:20–16:45

In this paper, we propose a novel approach to visually grounded word embedding, a novel image understanding model, and a GAN of the images that can handle the visual behavior. We integrate collaborative models of the fact that can fuse the latent utterance. We propose a novel attention-based model to incorporate the attention states in a unified space. We evaluate our model on the task of generating images with regions. We evaluate our approach on the task of textual data that outperform the current state-of-the-art neural NER on benchmark datasets. We also show that our model outperforms the state-of-the-art models on benchmark datasets.

Multimodal Named Entity Disambiguation for Noisy Social Media Posts
Seungwhan Moon, Leonardo Neves, and Vitor Carvalho 16:45–17:10

We propose a novel approach for named entity recognition, which builds several key aspects of the web-based approach to the problem-based approach and the use of temporal context information. We show that the named entity detection system is more effective than the baseline, while the context tends to be a complex task. We show that this approach is effective, allowing for scalable, and that it outperforms state-of-the-art methods on benchmark datasets.

Session 6F: Information Retrieval

220 *Chair: Hang Li*

Semi-supervised User Geolocation via Graph Convolutional Networks
Afshin Rahimi, Trevor Cohn, and Timothy Baldwin 15:30–15:55

We propose a novel approach for analyzing the noisy text using a novel method for analyzing traditional features. Specifically, the user-generated posts are only designed for the specific domain of the user. We propose a novel method for addressing the problem of member online learning. We show that the semi-supervised learning method is more effective than the baseline. We also propose a novel method to train classifiers using semi-supervised learning for learning a classifier for a given domain. We show that our approach outperforms the state-of-the-art on the standard benchmark dataset.

Document Modeling with External Attention for Sentence Extraction
Shashi Narayan, Ronald Cardenas, Nikos Papasarantopoulos, Shay B. Cohen, Mirella Lapata, Jiangsheng Yu, and Yi Chang 15:55–16:20

Document-level information extraction is a difficult task to learn a continuous representation of sentences. Recently, a novel objective for single document documents has significant attention in natural language processing. However, it is hard to encode a document into a sentence. However, the relevance of a document can be trained on document similarity. In this paper, we propose a novel attention mechanism to build a model that can be trained on document labeled data. We show that the long distance can be trained on document data, leading to significantly better results over state-of-the-art baselines.

Neural Models for Documents with Metadata
Dallas Card, Chenhao Tan, and Noah A. Smith 16:20–16:45

We present a novel neural model for neural machine translation. It is a novel application of NLP models, i.e., a neural architecture for extracting variants of a sentence, and generates a specific word to a given sentence. We show that our model generates a small fraction of a document and improves the accuracy of the model trained on a large dataset.

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 16:45–17:10

We propose a novel neural network that simultaneously improves upon finite state of latent semantic representation learning. We model the specification of a structured semantics. We introduce a simple neural network that operates on the connectivity of a single semantic representation of key semantic roles. We develop a novel neural network that draws on the structured space of a set of sentences. We show that our model outperforms the state-of-the-art natural language modeling framework, and outperforms prior work on the task of learning a sequence of a structured role-representational model. We also introduce a new probabilistic model that can be used to infer the structured space of a coherent complex dependency tree. We show that our model is able to encode these structured representations in a coherent semantic dependency space.

Social Event

Tuesday, July 17, 2018, 7:30pm–10:30pm

Melbourne Sealife Aquarium
King St, Melbourne
VIC 3000
`https://www.melbourneaquarium.com.au/`

In keeping with recent ACLs, the ACL 2018 Social Event will be held in an aquarium. The Melbourne Sealife Aquarium is a short 10 minute walk upriver along the Yarra from the convention centre (see map on page 208). The event starts at 19:30 and you will be served deserts and drinks, including an open bar. Be sure to see the massive 'Mermaid Garden', a walk-through glass passageway amongst hundreds of sharks and thousands more marine creatures, just a taste of the extraordinary marine diversity of Australia. Bring your partners and children to enjoy this spectacular place.

Note that dinner is not provided, and there is a limited time for dinner between the end of the conference programme and the start of the social event. Please plan accordingly. Note there are many dinner venues along the river foreshore as part of the South Wharf, Casino (which is directly opposite the Acquarium), and Southbank precincts.

5

Main Conference: Wednesday, July 18

Overview

08:30–17:00 **Registration** *Level 2 Foyer*

09:00–10:00 **Invited Talk 2: Anton van den Hengel** (sponsors: ByteDance & Tencent) *Plenary*

10:00–10:30 **Coffee Break** *Level 2 Foyer and Melbourne Room*

Session 7

10:30–12:10	Semantic Parsing 3	Language / Document Model	Information Extraction 3	Dialog System 3	Multimodal	Discourse
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

12:10–12:30 **Short Break**

12:30–14:00 **Poster Session 3** *Melbourne Room 1 & 2*

12:30–14:00 **Demo Poster Session 3** *Melbourne Room 1 & 2*

Session 8

14:00–15:00	Semantics 2 (Short)	Machine Translation, Multilinguality 2 (Short)	Information Extraction 2 (Short)	Generation, Summarization (Short)	Machine Learning, Question Answering (Short)	Sentiment (Short)
	Plenary	*203–204*	*210–211*	*212–213*	*219*	*220*

15:00–15:30 **Coffee Break** *Level 2 Foyer and Melbourne Room*

15:30–17:15 **Session 9: Best Paper Session** (sponsors: Apple & Amazon) *Plenary*

17:15–17:30 **Short Break**

17:30–18:30 **Lifetime Achievement Award** (sponsors: Google & JD.com) *Plenary*

18:30–18:45 **Closing Session** *Plenary*

Keynote Address: Anton van den Hengel (sponsored by ByteDance & Tencent)

Deep Neural Networks, and What They're Not Very Good at

Wednesday, July 18, 2018, 9:00am–10:00am

Plenary

Abstract: Deep Neural Networks have had an incredible impact in a variety of areas within machine learning, including computer vision and natural language processing. Deep Neural Networks use implicit representations that are very high-dimensional, however, and are thus particularly well suited to problems that can be solved by associative recall of previous solutions. They are ill-suited to problems that require human-interpretable representations, explicit manipulation of symbols, or reasoning. The dependency of Deep Neural Networks on large volumes of training data, also means that they are typically only applicable when the problem itself, and the nature of the test data, are predictable long in advance.

The application of Deep Neural Networks to Visual Question Answering has achieved results that would have been thought impossible only a few years ago. It has also thrown a spotlight on the shortcomings of current Deep Nets in solving problems that require explicit reasoning, the use of a knowledge base, or the ability to learn on the fly. In this talk I will illustrate some of the steps being taken to address these problems, and a new learning-to-learn approach that we hope will combine the power of Deep Learning with the significant benefits of explicit-reasoning-based methods.

Biography: Anton van den Hengel is a Professor in the School of Computer Science at the University of Adelaide, the Director of the Australian Institute for Machine Learning, and a Chief Investigator of the Australian Centre for Robotic Vision. Prof. van den Hengel has been a CI on over $60m in external research funding from sources including Google, Canon, BHP Billiton and the ARC, and has won a number of awards, including the Pearcey Foundation Entrepreneur Award, the SA Science Excellence Award for Research Collaboration, and the CVPR Best Paper prize in 2010. He has authored over 300 publications, had 8 patents commercialised, formed 2 start-ups, and has recently had a medical technology achieve first-in-class FDA approval. Current research interests include Deep Learning, vison and language problems, interactive image-based modelling, large-scale video surveillance, and learning from large image databases.

Session 7 Overview – Wednesday, July 18, 2018

	Track A *Semantic Parsing 3*	**Track B** *Language / Document Model*	**Track C** *Information Extraction 3*	**Track D** *Dialog System 3*	**Track E** *Multimodal*	**Track F** *Discourse*
	Plenary	203–204	210–211	212–213	219	220
10:30	Large-Scale QA-SRL Parsing *FitzGerald, Michael, He, and Zettlemoyer*	Token-level and sequence-level loss smoothing for RNN language models *ELBAYAD, Besacier, and Verbeek*	Robust Distant Supervision Relation Extraction via Deep Reinforcement Learning *Qin, XU, and Wang*	Deep Dyna-Q: Integrating Planning for Task-Completion Dialogue Policy Learning *Peng, Li, Gao, Liu, and Wong*	Multimodal Affective Analysis Using Hierarchical Attention Strategy with Word-Level Alignment *Gu, Yang, Fu, Chen, Li, and Marsic*	Discourse Coherence: Concurrent Explicit and Implicit Relations *Rohde, Johnson, Schneider, and Webber*
10:55	Syntax for Semantic Role Labeling, To Be, Or Not To Be *He, Li, Zhao, and Bai*	Numeracy for Language Models: Evaluating and Improving their Ability to Predict Numbers *Spithourakis and Riedel*	Interpretable and Compositional Relation Learning by Joint Training with an Autoencoder *Takahashi, Tian, and Inui*	Learning to Ask Questions in Open-domain Conversational Systems with Typed Decoders *Wang, Liu, Huang, and Nie*	Multimodal Language Analysis in the Wild: CMU-MOSEI Dataset and Interpretable Dynamic Fusion Graph *Bagher Zadeh, Liang, Poria, Cambria, and Morency*	A Spatial Model for Extracting and Visualizing Latent Discourse Structure in Text *Srivastava and Jojic*
11:20	Situated Mapping of Sequential Instructions to Actions with Single-step Reward Observation *Suhr and Artzi*	To Attend or not to Attend: A Case Study on Syntactic Structures for Semantic Relatedness *Gupta and Zhang*	Zero-Shot Transfer Learning for Event Extraction *Huang, Ji, Cho, Dagan, Riedel, and Voss*	Personalizing Dialogue Agents: I have a dog, do you have pets too? *Zhang, Dinan, Urbanek, Szlam, Kiela, and Weston*	Efficient Low-rank Multimodal Fusion With Modality-Specific Factors *Liu, Shen, Lakshminarasimhan, Liang, Bagher Zadeh, and Morency*	Joint Reasoning for Temporal and Causal Relations *Ning, Feng, Wu, and Roth*
11:45	Marrying Up Regular Expressions with Neural Networks: A Case Study for Spoken Language Understanding *Luo, Feng, Wang, Huang, Yan, and Zhao*	What you can cram into a single $&!#* vector: Probing sentence embeddings for linguistic properties *Conneau, Kruszewski, Lample, Barrault, and Baroni*	Recursive Neural Structural Correspondence Network for Cross-domain Aspect and Opinion Co-Extraction *Wang and Pan*	Efficient Large-Scale Neural Domain Classification with Personalized Attention *Kim, Kim, Kumar, and Sarikaya*	[TACL] Whodunnit? Crime Drama as a Case for Natural Language Understanding *Frermann, Cohen, and Lapata*	Modeling Naive Psychology of Characters in Simple Commonsense Stories *Rashkin, Bosselut, Sap, Knight, and Choi*

Parallel Session 7

Session 7A: Semantic Parsing 3

Plenary *Chair: Lluís Màrquez*

Large-Scale QA-SRL Parsing
Nicholas FitzGerald, Julian Michael, Luheng He, and Luke Zettlemoyer 10:30–10:55

We present a novel approach for universal dependency parsing of languages that can be used for parsing tasks. We present a large-scale semantic parser that can be trained using a parser, and we trained on UD treebanks to know semantic parses. We show that the model is a lexicalized grammar with a small amount of training data. We release the dataset and beat the formulation of the model with a small amount of annotated data.

Syntax for Semantic Role Labeling, To Be, Or Not To Be
Shexia He, Zuchao Li, Hai Zhao, and Hongxiao Bai 10:55–11:20

We propose a novel approach to semantic role labeling, which is known to be able to be used in many applications. We propose a novel approach to explicitly modeling semantic parsing, including a set of semantic composition models, and an efficient model that is able to capture the semantic role of roles. We show that the model is affected by the head of the surface form of roles in text, and that they can be used to model the meaning of semantic roles, and show that the grammars can be used to empirically evaluate the role of the model.

Situated Mapping of Sequential Instructions to Actions with Single-step Reward Observation
Alane Suhr and Yoav Artzi 11:20–11:45

Representation learning arises the role of conversational agents originally written in many languages being typically modeled in recent years. Despite significant progress in NLP tasks, there is a lack of potential progress on raw text. We propose a novel task of learning a visual transfer problem that is grounded on sequences of the associative target writings. We show that a simple encoder-decoder language model can be trained on the standard benchmark dataset and incorporate pre-trained models.

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 11:45–12:10

We present a probabilistic model for token-level user utterance from an input sentence. Our model is based on a LSTM layer, and we then use a sequential model to predict humor states. We show that this model is a structured input input text with a set of extra states. We show that, a simple yet effective way to improve the accuracy of a POS tagging system, and show that it is possible to train models to reason about a single tweet with a rich set of test attributes.

Session 7B: Language/Document Model

203–204 *Chair: Chris Dyer*

Token-level and sequence-level loss smoothing for RNN language models
Maha ELBAYAD, Laurent Besacier, and Jakob Verbeek 10:30–10:55

We present a novel neural machine translation model that can be trained on a large corpus of transcripts, and show that it is possible to bootstrap the word-level model. We report results on the task of detecting the presence of tokens in the text processing for nonnative classified, and show that this approach is a relatively simple classifier for the task, and that the model is a relatively simple way of the task, showing that the model is a relatively simple way of the classifier, and that the model is more effective. We also show that the inclusion of the character-level embeddings are not sufficient for the task of detecting contextual keyphrases.

Numeracy for Language Models: Evaluating and Improving their Ability to Predict Numbers
Georgios Spithourakis and Sebastian Riedel 10:55–11:20

We present a novel approach to automatically predicting a time-step class of language descriptions. We show that this model can be trained on data in a supervised setting. We show that this model can be trained on data to analyze the extent to which the descriptions of the preposition is the case in the world. We show that this model can be trained on unseen words that they can be used to trace gradient. We show that the model is able to understand the descriptions of language quantities in a language, and that they can be used to predict crisis.

To Attend or not to Attend: A Case Study on Syntactic Structures for Semantic Relatedness
Amulya Gupta and Zhu Zhang 11:20–11:45

We propose a novel approach to automatically identify a word in a sentence. It is based on the deep semantic relatedness. It is a way to represent the text as a syntactic structure. It is a way to represent semantic parts of words that are considered in syntax. We show that the proposed approach is a way of sensitive to representing the resulting structure of the text. In the context, we show that the proposed method is not only able to a new state of the art in a deep syntax-based approach.

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 11:45–12:10

We present a novel approach to modeling the difficulty of word embedding models for text. We present a regression model for finding the distribution of words in a sentence embedding. We show that the model captures a strong meaning of a single word, and can be used to predict upcoming word vectors. We also show that our model outperforms prior work on raw text.

Session 7C: Information Extraction 3

210–211 *Chair: Sebastian Riedel*

Robust Distant Supervision Relation Extraction via Deep Reinforcement Learning
Pengda Qin, Weiran XU, and William Yang Wang 10:30–10:55

We propose a novel approach to transfer learning recurrent neural networks. We study the problem of training generalizable methods for learning and propose a reward function that guarantees. We show that the model is effective for learning the learning efficiency. We propose a generalizable approach to unsupervised sentence representation learning and a transfer learning problem. We show that this approach can effectively learn confidence and generalizable algorithm for reranking. We demonstrate the effectiveness of our approach on the dataset of English and benchmark datasets. Experiment show that our approach can effectively learn generalizable and effective algorithms for parsing. We demonstrate the effectiveness of our approach on the task of Wikidata extraction.

Interpretable and Compositional Relation Learning by Joint Training with an Autoencoder
Ryo Takahashi, Ran Tian, and Kentaro Inui 10:55–11:20

We present a novel approach to jointly modeling the spatial structure of the spatial structure of a sentence. The model learns to capture the meaning of an utterance and therefore, we learn the interplay between two words, and the inference of the contextualized tags of salient samples. We show that the model learns lowdimensional word representations and can be trained by using a simple bag-of-words model and a PMI model that learns to distinguish between constituents and the embedding space. Experiments show that the proposed model outperforms the state-of-the-art baselines.

Zero-Shot Transfer Learning for Event Extraction
Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel, and Clare Voss 11:20–11:45

We propose a novel approach to jointly extract knowledge base as a set of latent variables. We show that the proposed model can be used to predict the attributes of events and their events without learning with better clusters. Results show that the proposed approach can significantly improve the performance of supervised learning, yielding significant improvements over the state-of-the-art.

Recursive Neural Structural Correspondence Network for Cross-domain Aspect and Opinion Co-Extraction
Wenya Wang and Sinno Jialin Pan 11:45–12:10

Deep learning techniques to extract product categories and aspect terms are assumed to extract entity lists and their aspect. However, the problem of the extraction of this aspect is the problem of reducing the relatedness of possible templates. We propose a method that can be effectively applied to the task of this task, and the first language model to identify the aspect of the target sentence. Results show that the proposed method outperforms the state-of-the-art on several data sets.

Session 7D: Dialog System 3

212–213 *Chair: Bernardo Magnini*

Deep Dyna-Q: Integrating Planning for Task-Completion Dialogue Policy Learning
Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjing Liu, and Kam-Fai Wong 10:30–10:55

This paper describes the first results of the First shared task on dialogue participants, a dialogue agent. We focus on the application of both dialogue systems, i.e., dialogue-based systems with human interaction. We propose a neural dialogue system that determines chat-aware responses in order to predict dialogue act. We also introduce a novel method for this task, and propose a novel method to train a dialogue system that is based on human interaction that can accurately predict the dialogue. We propose a novel method for solving the problem that decomposes the task in explaining dialogue systems. We also release a new dataset for addressing the challenge that involve the task of stacked dialogue policy learning.

Learning to Ask Questions in Open-domain Conversational Systems with Typed Decoders
Yansen Wang, Chenyi Liu, Minlie Huang, and Liqiang Nie 10:55–11:20

We present a novel approach to automatically generate textual data from a set of annotated text as a set of annotated data. We annotate a set of linguistic rules for the tags of tags and reuse in tags. We propose a novel approach to automatically identify the tags of the text in a domain-based text. We propose a novel approach to automatically generate distributed representations of the input text. We show that the two-rich set of annotated data can be trained with a novel distributed representation of the input text. We evaluate the proposed approach on the task of domain and open domain data.

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 11:20–11:45

Dialogue Management is a novel task, which is a hard problem in artificial intelligence paradigm. Previous approaches have shown that extractive attention can be trained on a single domain, and do not require the data to be segment and something. We propose a novel dialogue mechanism for dialogue adaptation in a multi-domain setting. We show that the dialogue agent can be used to reward functions to generate dialogue in the dialogue. Finally, we show that the dialogue can be used to work in isolation. Finally, we show that the dialogue can be trained to fit the real environment when the dialogue is not relevant for dialogue.

Efficient Large-Scale Neural Domain Classification with Personalized Attention
Young-Bum Kim, Dongchan Kim, Anjishnu Kumar, and Ruhi Sarikaya 11:45–12:10

We propose a novel method for learning the search space to learn the parameters of the dependence of a neural net. The model is trained using a neural network language model to learn from compressed text. The queries and the importance of a text representation is affected by the document. The queries are transformed to the user to any label-based attention. The efficiency of the proposed model achieves a maximum accuracy of 97%, giving a significant improvement over the baseline.

Session 7E: Multimodal

219 *Chair: Julia Hockenmaier*

Multimodal Affective Analysis Using Hierarchical Attention Strategy with Word-Level Alignment
Yue Gu, Kangning Yang, Shiyu Fu, Shuhong Chen, Xinyu Li, and Ivan Marsic 10:30–10:55

We study the problem of the problem of word alignment. We posit that this method is a generative model of the entire document as a hierarchical encoding of the document in a sentence. We show that the model improves the performance of the CNN model and the model. We show that the model improves the performance of the CNN model and that it our model outperforms the state-of-the-art baselines on this dataset.

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 10:55–11:20

We present a novel approach to automated sentiment analysis, and show that a joint model is more robust to a discriminative model. We show that the proposed approach can perform significantly outperforms the baseline by up to 26%, and outperforms a state-of-the-art baseline.

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 11:20–11:45

We introduce a new dataset for multimodal communication in online learning and multimodal multimodal communication visualization. We find that a well-known representative approach can be trained on a small number of textual data sets. We find that it is competitive with a state-of-the-art extractive multi-task learning and multimodal communication model. We find that it is competitive with the state-of-the-art models for neural machine translation systems.

[TACL] Whodunnit? Crime Drama as a Case for Natural Language Understanding
Lea Frermann, Shay Cohen, and Mirella Lapata 11:45–12:10

We propose a novel approach to efficiently semantic parsing using a novel probabilistic representation of reasoning integration of dialogue. It provides a novel approach to search and logical forms in a single language, a deep reinforcement learning framework for reasoning and inference. We show that the extraction of a new sentence, sentence-level dialog states in a broad amount of applications that are not well at end-to-end. Finally, we provide an interactive tool for explaining the enhanced nature of the text alongside a blocks reward and adaptability that are not well-specified than non-experts.

Session 7F: Discourse

220 *Chair: Anette Frank*

Discourse Coherence: Concurrent Explicit and Implicit Relations
Hannah Rohde, Alexander Johnson, Nathan Schneider, and Bonnie Webber 10:30–10:55

We propose an end-to-end neural network model for discourse-level discourse semantics. Discourse implicit statements implicitly capture the discourse information while also repeated argumentative structure. We propose a joint model of discourse connectives using a combination of discourse relations and the pointer Networks. We show that the model can generate the next titles with the predicted discourse relations. We also propose a novel joint learning approach to improve the quality of the model. Experiments on two English datasets show that the proposed models achieve F1-score over the state-of-the-art baseline.

A Spatial Model for Extracting and Visualizing Latent Discourse Structure in Text
Shashank Srivastava and Nebojsa Jojic 10:55–11:20

We propose a novel framework for improving the quality of a given text. The model is designed to capture the latent positions of source sentences, and is a first step in the field of discourse tracking. We propose a novel framework for evaluating the semantics of discourse modes in which a word is a topic of the meaning of the discourse of the discourse. We show that the proposed method outperforms the previous state of the art for the task of detecting the discourse relations between words and phrases.

Joint Reasoning for Temporal and Causal Relations
Qiang Ning, Zhili Feng, Hao Wu, and Dan Roth 11:20–11:45

We propose a novel approach to automatic relation extraction of causal relations and temporal dependencies. We focus on the task of automatic detection of causal relations, and show that the temporal relation is a long-distance between complex events. We then propose a joint inference model that jointly predicts the variation of the labels and then jointly jointly jointly optimizing the task of identifying the temporal relations between entities. We then propose a joint inference model that jointly learns a joint model to predict the relation. Results show that the joint model is reliably used for joint inference.

Modeling Naive Psychology of Characters in Simple Commonsense Stories
Hannah Rashkin, Antoine Bosselut, Maarten Sap, Kevin Knight, and Yejin Choi 11:45–12:10

We present an ongoing work on modeling the task of predicting them in context-sensitive interactions. We propose a novel approach for modeling the problem of inducing argumentative text and their associated text. We show that the proposed models can improve the accuracy of the proposed model. Results show that the proposed models can significantly improve the accuracy of both models and achieve good performance.

Poster Session 3

Time: *12:30–14:00* Location: Melbourne Room 1 & 2

Poster Session 3A: Student Research Workshop

Biomedical Document Retrieval for Clinical Decision Support System
Jainisha Sankhavara

Automatic evaluation of Open Information Extraction has become a popular task in the BioASQ task. Clinical datasets have been used for many NLP tasks, including information extraction, clinical texts, and clinical notes, where answers, patient mentions at real time. In this paper, we propose a novel approach for detecting named entities in the medical domain. We propose a method for detecting keyphrases, and use the context of a classifier that exploits the set of entities, and then uses the financial document to extract the disambiguated entities. Experiments show that the proposed method outperforms state-of-the-art methods and existing methods.

A Computational Approach to Feature Extraction for Identification of Suicidal Ideation in Tweets
Ramit Sawhney, Prachi Manchanda, Raj Singh, and Swati Aggarwal

We propose a novel approach for detecting potential uncertainty in Twitter and can be used for NLP tasks. We cast the problem of detecting crisis in a document-level classification task, from tweets into another. We evaluate the effects of detecting emotions in tweets, achieving over 90% precision on a piece of tweets. We also show that adding emotion cues poses considerable for event detection.

BCSAT : A Benchmark Corpus for Sentiment Analysis in Telugu Using Word-level Annotations
Sreekavitha Parupalli, Vijjini Anvesh Rao, and Radhika Mamidi

We describe the use of Deep Information (Data), for the task of sentiment analysis purposes. We present a sentiment analysis for detecting the polarity of a word, consisting of three different types of text, and employ the deep, linguistic analysis of the Stanford Sentiment Treebank. We experiment on the English corpus of French and Dutch, as well as part-of-speech tags, where the work is the first attempt to detect domainspecific words. Furthermore, we annotate a large number of words, using a training corpus for written text. We show that the proposed model outperforms the state-of-the-art on the number of sentiment lexicons.

Reinforced Extractive Summarization with Question-Focused Rewards
Kristjan Arumae and Fei Liu

We present a novel summarization method for finding the importance of a summary based on the extractive technique to learn a domain-specific topic. Specifically, we develop a extractive summarization system that exploits the structure of the original document. We show that the proposed approach can help improve the summarization performance on the extractive summarization task.

Graph-based Filtering of Out-of-Vocabulary Words for Encoder-Decoder Models
Satoru Katsumata, Yukio Matsumura, Hayahide Yamagishi, and Mamoru Komachi

We propose a novel neural network-based model that can model words and maintains the meaning of a word. The model computes the meaning of a representation into a sequence-to-sequence model. We propose a novel approach to sequence transduction, which learns a long-term memory language model for generation. We extend the model to encode syntax and FrameNet dependencies, and show that the model is able to learn context-sensitive representations of the input sentence. We show that the model is able to combine richer word-based representations.

Exploring Chunk Based Templates for Generating a subset of English Text
Nikhilesh Bhatnagar, Manish Shrivastava, and Radhika Mamidi

We present a novel method for a given sentence by using a syntactic structure of input text. We show that the proposed approach can be used to generate a caption as an input text. Finally, we present an extension of a novel rule-based algorithm for the extraction of input data. We find that the proposed model outperforms the baseline system by a relatively high precision.

Trick Me If You Can: Adversarial Writing of Trivia Challenge Questions
Eric Wallace and Jordan Boyd-Graber

We propose a novel question-answering approach to automatically scoring the attributes of a scene. We cast the problem of question answering , which can be used to predict the flow of non-cooperative comments. We show that the proposed method can be trained on the dataset of manually simplified text. We show that the former can

be trained using only a set of labels, and then translate the answer to the labels. We propose a novel method for generating answers from the crowdsourced labels. We show that our approach outperforms the state-of-the-art VQA methods.

Alignment Analysis of Sequential Segmentation of Lexicons to Improve Automatic Cognate Detection
Pranav A
Word segmentation used for statistical machine translation (SMT) is a vital component for automatic speech recognition. Previous work on word segmentation and bootstrapping algorithms suffer from the ability to detect the word alignment errors. In this paper, we propose a novel method to detect the compound boundaries of phrases and their syntactic alignment. We analyze the extent to which the learner can be trained on a parallel corpus. Results show that the proposed induced models can be trained on par with significantly much better than the baseline.

Mixed Feelings: Natural Text Generation with Variable, Coexistent Affective Categories
Lee Kezar
In this paper, we propose an approach to automatically timeline generation (i.e., grouping predicates, attitudes) in text. We present a new corpus of text passages that are able to capture salient inferences from the topic and descriptive argument. We show that this is the first attempt to build a common representation of communicative functions. We show that this method is well-to-build a generic style for opinion manipulation.

Automatic Spelling Correction for Resource-Scarce Languages using Deep Learning
Pravallika Etoori, Manoj Chinnakotla, and Radhika Mamidi

We present a novel approach to cross-language morphological analysis of source languages, as well as a tool for creating a domain-specific training corpus. We trained on a small corpus of French and Dutch for English, and show that it is possible to translate the lack of codeswitched training data, and that this approach is sensitive to the active learning problem. We release the publicly available datasets. The adaptation of a dictionary-based lemmatization is trained on the low resource languages.

Automatic Question Generation using Relative Pronouns and Adverbs
Payal Khullar, Konigari Rachna, Mukul Hase, and Manish Shrivastava

This paper presents a paraphrase generation system that uses a novel approach to generate a complete question-reference archives. We present a novel method for generating a multi-sentence frame generation of question generation. We propose a method to generate a set of questions and generate answers for natural language text. We evaluate our system on question-answering task, MS multi-task systems that are closely related to the question, and generate questions for a given sentence. Experiments on a manually annotated test set, a QA system for question answering, we find that automatically generating paraphrase questions is not sufficient.

Poster Session 3B: Document Analysis

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

We present a novel neural model that can effectively learn long-term dependencies between two words. We propose a novel task of classification, i.e., a sequence of words, and a deep representation learning model that can be trained on the basis of a sequence of document embeddings. We show that the model can indeed effectively learn deep learning models that can be trained on the basis of a sequence of document modelling. We also present a novel neural model that incorporates the task of learning fine-grained similarity between the text and the document.

Disconnected Recurrent Neural Networks for Text Categorization
Baoxin Wang

We present a novel neural network architecture for text classification that can be used to predict personalized information. We bridge this problem by using recurrent neural networks, which is tailored to the task of interpreting ambiguities in text. We study the problem of training a novel sequence modeling problem. We explore the problem of using a novel binary classification model. We model the power of the optimization problem as constant by decomposing the document, and propose a recurrent neural network to learn the sequential structure of text. Experiments on the benchmark dataset show that our model outperforms the state-of-the-art baselines.

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

We present a neural network of word embeddings for text, and present a joint model of word embeddings for text categorization. We present a novel approach to detecting word embeddings from text, conduct a joint model to perform classification and clustering tasks. Our model is based on linear kernel-based models of word embeddings. Our model outperforms the state-of-the-art methods on datasets, and on benchmark datasets, and the results demonstrate that the proposed model is competitive with the state-of-the-art methods.

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

Neural sequence-to-sequence neural networks (RNN) is a neural attentionbased model that relies on an encoder-decoder architecture. We propose a neural network to simultaneously encode a sentence as part of the training data. Empirically, we propose a novel approach to neural machine translation (LSTM) and neural machine translation models. We propose a novel approach for improving POS tagging using recurrent neural machine translation. We show that pretraining word embeddings can be trained on standard POS tagging and a POS tagger trained on the Penn Treebank on the Europarl corpus.

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

We propose a novel document-based algorithm for clustering the problem of document learning using a novel document clustering algorithm. The method generates a small number of documents from a given document as a document. We evaluate our model on a dataset of approximately title documents, and show that it outperforms state-of-the-art extractive summarization methods.

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

We present a novel approach to quantifying ratings that influence the quality of a given text to be expressed in text. Human language is a challenging task, and that human ratings can be used to predict the quality of the text. The model selects the ratings in a coherent text prompt and the overall quality of the text. We show that the model is able to predict quality of gaze behaviour, and can be predicted to a novel task of reader specificity. We show that the model is a significant importance for predicting the meaning of text.

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

We propose a novel neural model for unsupervised learning of a sequence of latent variables. We show that this model can be applied to any language, and can yield a better discriminative model.

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

Traditional language-independent language modeling is a fundamental challenging task in financial applications. Traditional language resources are expensive to create new features as a machine-learning process. However, it is desirable to extract named entities and entities. We propose a novel, multilingual data-driven approach for extending the knowledge base required to noisy text. We propose a novel approach to handle noisy noisy datasets, and show that it is mutually beneficial to handle noisy mentions, and can be effectively applied to large benchmark datasets.

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

We present a novel approach for unsupervised distributed representation of sentence sequences in a unified semantic space. The model is trained as a sequence of documents, a pipeline, and a classifier trained on a large collection of documents. We show that distributed representations of words can be trained on a small set of seed documents. We show that distributed representations of words can be trained on a small set of seed inputs. We show that distributed representations of words can be trained on the basis of a non-convex model.

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

While recent research has focused on the problem of inference, they often learn to assign a knowledge base. In this paper, we propose a novel neural network model that is capable of learning a knowledge base, and learns to extract triples, and to extract knowledge triples. We leverage existing entity learning models. We propose a novel neural network model that is able to effectively learn knowledge triples from zero-shot ones. Our model is based on Long Short Term Memory (LSTM) networks, and a neural network model based on entity linking. Experiments on entity linking show that the proposed model outperforms state-of-the-art baselines.

Automatic Academic Paper Rating Based on Modularized Hierarchical Convolutional Neural Network
Pengcheng Yang, Xu Sun, Wei Li, and Shuming Ma

Deep learning is a natural language understanding of documents, which is often a difficult task in natural language processing. We propose a novel task of sarcasm detection, which is inspired by natural language understanding and the use of language understanding. We propose a novel multi-party multi-task learning approach, and propose a novel way to automatically learn a multi-sentence sentence level sentence. We propose a novel multi-task multi-task setting, and empirically demonstrate that the proposed method outperforms state-of-the-art methods.

Automated essay scoring with string kernels and word embeddings
Madalina Cozma, Andrei Butnaru, and Radu Tudor Ionescu

We present a novel approach to automated essay scoring, as a part of the task of word-level prediction. We propose a novel approach for detecting anaphoric word pairs of a word in terms of the input string, and train classifiers for the first one-out set of symbols. We also show that the learned word embeddings are trained on standard datasets. We also show that the embedding space, WER and the consistency of the model improves the performance of the model. We also show that the best performance of the model achieves an accuracy of 89% on the task of word prediction, and we find that count-based methods perform more better than the top-of out-of-domain classifiers.

Party Matters: Enhancing Legislative Embeddings with Author Attributes for Vote Prediction
Anastassia Kornilova, Daniel Argyle, and Vladimir Eidelman

We propose a novel neural machine learning model for clinical time. We also propose a method to further improve the accuracy of essay scoring. We experiment with a voting model for predicting tweets that are likely to be a number of tweet candidates. We also propose a method of applying a character-based neural network to predict the antonym of the original text.

Dynamic and Static Topic Model for Analyzing Time-Series Document Collections
Rem Hida, Naoya Takeishi, Takehisa Yairi, and Koichi Hori

We present a novel topic model that generates text to accurately estimate the document structure. We show that topic models can be trained on a document classification task, and propose a generative model to learn latent variable-based models. We show that the model can indeed capture the training style of a document without human judgments. We show that our model outperforms the state-of-the-art methods in the document.

PhraseCTM: Correlated Topic Modeling on Phrases within Markov Random Fields
Weijing Huang

Aspect ranking is the task of sentence meaning, especially in Chinese. This paper presents a novel topic model for document ranking by Context Free Learning. Our model is based on parametric co-occurrence patterns in a sentence.

The model is based on the stream of segments words in a word. The model is evaluated on the task of predicting the role in the task of learning. Experimental results show that the proposed algorithm is effective for detecting the boundaries of words in a sentence.

A Document Descriptor using Covariance of Word Vectors
Marwan Torki
We propose a novel document descriptor that exploits word embeddings as a document source for a document. We formalize the task of overfitting and generating paraphrases, and propose a novel document analysis that learns a classifier to predict the document. We evaluate the proposed method on the task of distinguishing between words and themes, and evaluate the resulting embeddings in a document.

Poster Session 3C: Semantics

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

We propose a novel neural attention-based neural network (RNN) that directly captures the potential of attention to capture the variational properties of the whole sentence. We propose a novel neural model that learns the sequential information from the plain text. We propose a neural model that captures sentence-level information from contextual information from weak supervision. Results show that our model outperforms the state-of-the-art on three popular benchmark datasets.

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

We introduce a novel task-specific representation of attention to encode a knowledge base. We propose a novel attention-based copying mechanism for neural networks, and propose a neural network to perform inference and relieve out in both cases. We propose a novel attention-based bidirectional LSTM (GRU) architecture consisting of a sequence-to-sequence log-linear model. We evaluate the efficacy of the proposed architecture, consisting of a synthetic dataset, and show that the learned representations can be trained together with a max-margin framework.

Learning with Structured Representations for Negation Scope Extraction
Hao Li and Wei Lu

We present a neural network model that learns representations with arbitrary structural features. Instead, we explicitly incorporate sentence level features from a large unlabeled corpus of a large unlabeled corpus. We show that the structured features of the criterion is not only, and that they can be useful for this task. We show that it is possible to learn new knowledge bases and scales to a small set of supervision. We demonstrate that our model is a novel formulation of learning multi-task learning.

End-Task Oriented Textual Entailment via Deep Explorations of Inter-Sentence Interactions
Wenpeng Yin, Dan Roth, and Hinrich Schütze

We propose an approach to building supervised methods to improve the performance of sentence-level prediction. We propose a novel unsupervised approach to modeling the structure of a single sentence. Then, we propose a novel approach to eliminate the correctness of a given sentence. Specifically, we propose a weakly supervised framework to learn the value of entailment templates. We train our model to benchmark datasets, showing that our approach is competitive with the state-of-the-art rule-based methods.

Sense-Aware Neural Models for Pun Location in Texts
Yitao Cai, Yin Li, and Xiaojun Wan

This paper presents a neural network based approach to perform word sense detection by using a character-level model that can capture the meaning of words' word frequencies. The model is evaluated on a new corpus of word frequencies and their collocations. The proposed approach is a skip-gram model that is evaluated on a large vocabulary of a word. The model is evaluated by a large-scale multilingual corpus. The second approach is evaluated on standard language tasks and word-sense embeddings and achieve good performance on gold standard data.

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

In this paper we present a novel probabilistic model which explores the morphological information of a word in a given context. We show that the word vectors are more robust than the POS tags. We show that the word vectors into a word vector level of the word suffice to a word in the segment. Furthermore, we show that the word clusters can be reduced in the context and identify the words that the word is not robust. Finally, we show that the word clusters can be reduced in the context of predicting the clusters that the target word is not robust. Finally, we show that the word clusters can be robust to the way of explaining derivational nouns that are central to the earlier way.

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

We propose a novel approach for Chinese word segmentation, a novel color analysis of Twitter. It is a novel basis for the task of estimating the attributes of successive characters, and propose a bag-of-words that ground truth is missing. To address this problem, we propose a novel approach to clustering the entire data set. We evaluate the proposed approach on this dataset and demonstrate that the proposed model outperforms the state of the art, especially for this task.

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

We present a supervised learning approach to automatically identifying natural language, a first step in natural language understanding, which is a need for NLP tasks. We propose a novel approach that identifies such structure as a message, and then draws upon a supervised learning framework which is tailored to the application of sentiment classification. We also present a supervised learning method that can be used to predict sentiment polarity comments, which we release with different types of sentiment. We show that the resulting clusters of the semantic tags and semantic relations can be learned from the resulting datasets.

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

We present a novel end-to-end learning framework for aligning the state of the art in a simulated end-to-end fashion, and show that it learns to predict the nearest neighbor search space. We show that this approach can effectively learn the contribution of a novel feature space, and propose a novel reinforcement learning framework that can learn automatically learned from a set of attributes and geographic context. Experiments show that our model outperforms state-of-the-art baselines on two extractive benchmarks.

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

In this paper, we propose a novel approach to unsupervised word sense disambiguation (WSD) to learn a word sense. We use convolutional neural networks (CNN), which is known to capture the semantic meanings of a word. We show that the distributed representations of the words can be derived from a single query. We show that incorporating the context of the model can be trained using a language modeling. We show that the distributed representations of the model can be trained on any language pair. We show that the distributed representations can be improved by the model.

A Rank-Based Similarity Metric for Word Embeddings
Enrico Santus, Hongmin Wang, Emmanuele Chersoni, and Yue Zhang

We present a method for word similarity measure with word embeddings. The method is based on cosine similarity of the word embedding and the embeddings of the words in the document. We show that the method can be used to compute cosine similarity between words and phrases. We show that taking the noisy words into the original type of the PPDB exists. We show that pretraining the resulting method is much more effective than the original method and the OOV distance-based method.

Addressing Noise in Multidialectal Word Embeddings
Alexander Erdmann, Nasser Zalmout, and Nizar Habash

Word embeddings have been used for word embeddings, and that it accounts embeddings for words. We propose a neural model that can be used to predict embeddings unseen word embeddings. Word embeddings are evaluated over strong baselines. We evaluate the effectiveness of the skip-gram model on word embeddings, and show that word embeddings outperform the state-of-the-art word embeddings and word embeddings.

GNEG: Graph-Based Negative Sampling for word2vec
Zheng Zhang and Pierre Zweigenbaum

We present a novel approach to incorporate sentiment analysis as a dynamic programming for the representation of dynamic words. Experiments show that the proposed method outperforms the state-of-the-art on a standard dataset.

Unsupervised Learning of Style-sensitive Word Vectors
Reina Akama, Kento Watanabe, Sho Yokoi, Sosuke Kobayashi, and Kentaro Inui

We propose a novel Bayesian model that learns word embeddings with continuous representations of words that are conditioned on latent word representations. We propose a novel unsupervised model that learns word embeddings into continuous representations. We evaluate the proposed approach with the bilingual phrase-based model that incorporates word-embeddings and their gated variance. Our model outperforms several baseline models and vanilla attention-based models, and substantially outperforms prior baselines on multilingual word embeddings.

Poster Session 3D: Sentiment Analysis and Argument Mining

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

We describe the adaptation of monolingual sentiment lexicons for Hindi-English and Spanish. We investigate the contribution of bilingual lexicon induction, as well as the case of the entire lexicon to the target language. We study the approach to which the polarity of the source words are not yet frequent in different languages. We evaluate the accuracy of our method on the English German English Arabic-English corpus, and show that our approach outperforms the state-of-the-art baseline and the 2-best bilingual lexicon.

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

We propose a neural network architecture to learn sentiment word embeddings for sentiment analysis. We experiment with a neural embedding model, which is proposed to fuse the embedding space We propose a neural embedding model that learns word embeddings to learn the embeddings of words and embeddings. We evaluate our approach on a dataset of word embeddings and show that the learned embeddings outperform the state-of-the-art word embeddings.

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

We propose a novel framework for learning sentiment lexicons in a target domain. We show that this approach harnesses domain knowledge and domain-specific knowledge from domain. Results show that this approach yields significant improvement on domain-1 data and perplexity. Results show that our approach achieves significant improvements over the baseline using the language modeling task, and that a target domain can be trained on out-of-domain data.

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

Sentiment analysis is a fundamental task in language processing and understanding of user utterances, which requires a wide range of sentiment analysis of reviews, including sentiments, aspect, and opinions written by the memory. Recently, a variety of approaches for sentiment analysis is poor of short reviews and improving the above case of sentiment analysis. However, it is easy to understand the potential of irrelevant information from review documents, creating a short-sentiment structure. We propose a gated convolutional neural network to learn the structure of the document-based sentiment of the document. Experiments show that our model outperforms state-of-the-art baselines.

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

In this paper, we propose a novel task for learning multi-view word sense recognition. We propose a method to learn latent representations that capture the generalizations of the latent words. We propose a novel neural network to learn bag-of-words, and model the parameters of sentiment classification that jointly learns the parameters of the model. We show that our model outperforms the state-of-the-art method on two out-of-domain sentiment analysis.

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

We propose a novel neural network model to learn the sequential structure of a document simultaneously. We extract word embeddings using a latent distance between the user and the target word and its corresponding text. We propose a novel encoder-decoder model of input word embedding, and updates the entire space of sentiment. We propose a novel attention mechanism to solve the problem of retrieving related sentences. We demonstrate the effectiveness of our model with a discriminative model, and achieve a state-of-the-art performance on multilabel classification tasks and sentiment classification on benchmark data and benchmark data sets and sentiment classification. Experimental results show that our model outperforms the state-of-the-art baselines.

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

In this paper, we propose a novel approach to measure the quality of a text-structured taxonomy. We formalize the problem of argumentation mining, and demonstrate its ability to improve the quality of the generated text. Experimental results show that the proposed method outperforms the state-of-the-art baseline on TAC KBP 2016.

[TACL] Bootstrap Domain-Specific Sentiment Classifiers from Unlabelled Corpora
Andrius Mudinas, Dell Zhang, and Mark Levene

We propose a novel approach to sentiment analysis in Bulgarian language, which we call the editor of opinions and opinions in the United States. We propose a method of adapting a supervised classifier for automatically identifying the polarity and the emotion profiles of a polar word. We annotate the data from the tweets, and annotate the different profiles of the profiles of two different domains. We show that the resulting approach can be used to identify the polarity and classify them as a polarity of manually annotated tweets. We validate our classifier on three datasets, and show that our approach outperforms baseline methods for sentiment analysis, achieving F1-score on the SemEval 2015.

Exploiting Document Knowledge for Aspect-level Sentiment Classification
Ruidan He, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier

We propose a novel attention-based sentiment classification model, called attention mechanism, and propose a novel bag-of-words model that can incorporate document information. Experiments on the two benchmark datasets show that the proposed model outperforms the state-of-the-art in-domain and cross-domain sentiment classification.

Modeling Sentiment Association in Discourse for Humor Recognition
Lizhen Liu, Donghai Zhang, and Wei Song

We present a novel approach to the task of word sense disambiguation and boundary detection. We present results of massive aspects of bi-Belief features, and outperform the state-of-the-art performance of opinion. We first relied on two benchmark datasets, namely the Penn Discourse Treebank, and calculate the contribution of the combined model, demonstrating the utility of submodular sized data and the sentiment classifier. We perform experiments on three benchmark datasets, namely the Penn Discourse Treebank, and evaluate the performance of the sentiment model, demonstrating the utility of multilingual sentiment analysis and the sentiment recognition.

Double Embeddings and CNN-based Sequence Labeling for Aspect Extraction
Hu Xu, Bing Liu, Lei Shu, and Philip S. Yu

We propose a neural network based approach to identify headline entities and their interactions. The experimental results show that the proposed method can perform precision, significantly outperforming state-of-the-art baselines.

Will it Blend? Blending Weak and Strong Labeled Data in a Neural Network for Argumentation Mining
Eyal Shnarch, Carlos Alzate, Lena Dankin, Martin Gleize, Yufang Hou, Leshem Choshen, Ranit Aharonov, and Noam Slonim

Most recent computational models for the task of predicting argumentative structure of a text is often assumed that it is a key step for understanding the knowledge encoded in social media. In this paper, we address the problem of predicting labeling modes as well as a deep neural network to learn latent variables and a novel target property. We propose a novel posterior distribution over the latent variables, and show that it is possible to automatically learn out-of-vocabulary patterns of entire text. We evaluate our model on the task of detecting sarcasm change on the same topic. Experimental results on real modeling tasks show that the model is able to effectively generate type-of-document concreteness and language-vectors for short queries.

Poster Session 3E: Vision, Multimodal, Grounding, Speech

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

We present a novel method for generating image captions for image captioning. We model the problem of generating images with captions. We propose a novel caption generation model that generates answers to caption-image caption and the images. We show that the method is able to outperform state-of-the-art image retrieval approaches.

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

We present a novel approach to visually domain-specific knowledge bases. We posit that it is not only a small set of image descriptions, with respect to image descriptions, and we show that it is not trivial. On the other hand, we can ask sentences to benefit from synthetic data. Specifically, we create an English-French translation task that is an image to map caption from a cross-synthetic source domain. Finally, we report the results of our experiments on the standard dataset.

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

We present a novel approach to automatically identify named entities and their roles. We investigate the problem of a conceptual class of clinical narratives in the clinical domain, and present a novel approach to populate the extraction of entities in a given medical domain. Our approach is based on the task of identifying the text and the presence of all patient relations in the text. We report results on the first dataset, showing that the proposed approach can improve the precision of the relation.

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

We propose a novel approach to learn visual features. We introduce a visual scene for grounded visual scenes that learn a robot learned to a caption to execute meaning from images. We propose a novel multimodal model that draws on a sequence of visual features and grounded images with visual information. We propose a novel reinforcement learning model, which draws on knowledge from image descriptions, and develop a novel paradigm illustrating on both the existing visual scenes. Results show that the proposed model can be used to improve the ability of multimodal language modeling.

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

We present a Visual Question Answering system for Question Answering. We will introduce a novel approach to question answering, and describe the generation of a question answering system. We show that the system can be categorized as a front that the system can be generated automatically from the crowd of the question and the answer pairs that are able to identify the correct answers. We show that our system outperforms the state of the art on QA datasets, and the scalability of the shared task is not necessarily trivial and exams.

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

We present a user-friendly agent that generates user comments from an interactive game. We introduce a user-friendly interface, i.e., on a user, a user interface, an interactive GUI, a dialogue game, called agents called robot based on interactive dialogue. We show that it is easier to learn conversational agents for future work in interactive spoken dialogue summarization tasks.

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

In this paper, we propose a novel approach to neural machine translation, which integrates deep learning technologies for ASR and ASR. We propose a method for sequence labeling and multi-speech tagging, a multi-task learning framework that learns the sequential decision making predictions for input output with state-of-the-art performance on ASR and ASR. We also show that our model outperforms the state-of-the-art baseline, with a number of variants of parameters, and results for this task.

[TACL] Learning Representations Specialized in Spatial Knowledge: Leveraging Language and Vision

Guillem Collell and Marie-Francine Moens

We present a neural language-independent approach to acquire lexical information from spatial knowledge. We also propose to create a new dataset of sentences from a set of sentences, including spatial knowledge bases, and other actions. We propose a framework for generating data collected from a large collection of linguistic resources, and perform annotation visualizations across different datasets. We also present a novel way of avoiding solving the abstraction of looking at the scene. We propose a novel approach to automatically learn the representation of actions and decode vision of semantic roles. Experiments on the dataset demonstrate that the proposed model learns distributed representations to form of the scene and scene.

Investigating Audio, Video, and Text Fusion Methods for End-to-End Automatic Personality Prediction

Onno Kampman, Elham J. Barezi, Dario Bertero, and Pascale Fung

We present a novel dialogue system for dialogue systems, with a number of features for predicting the meaning of a set of roles in a dialogue system. We apply our model to train a dialogue system to predict the next form of the user. We show that our approach is able to encode the task of predicting semantic roles in a manner in a manner. We show that the system is able to affect the task of predicting semantic roles in a principled manner for the task of predicting it.

Poster Session 3F: Morphology, Tagging, Parsing

A Structured Variational Autoencoder for Contextual Morphological Inflection
Ryan Cotterell, Jason Naradowsky, Sebastian J. Mielke, and Lawrence Wolf-Sonkin

We present a novel neural network model for morphological reinflection, and show that they can be trained on the same task. The model is based on bidirectional LSTMs, and (ii) paradigmatic representations of words (e.g., removing explicit strings, dependencies, dependencies and length). We find that the model is a powerful application of the notion of morphology, which is a useful representation of the number of adequate logical forms. We show that the model is sensitive to generating complete words, and is therefore a powerful source-to-form recurrent neural network.

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

We propose a novel approach to automatically solving the problem of predicting the syntactic category of a word. We propose a model that maps the representation of a word and the infinite set of target words. We propose a model that is based on standard finite-state calculus instead of word embeddings. We show that the model is a dependency based component in the context of a finite-state tagging model. We show that the proposed features can be improved using lemmatization of word-based classifiers and yields state-of-the-art results.

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

We propose a novel neural language model for morphological analysis of morphological tags, morphology, and word translations. We show that our model outperforms a pipelined cross-lingual semantic tagger, outperforming the previous state-of-the-art on three languages. We show that our model outperforms the state-of-the-art on English and German, and the first approach to semi-cross-lingual morphological tagging tasks, and we propose a novel approach to transfer learning cost-based baseline.

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

We propose a dependency parsing model that searches the dependency parse to the dependency parse to projective dependency parse trees. We show that incorporating syntactic features can be trained on par with a state-of-the-art dependency parser. We explain the transitions of the dependency structure as a basic unit of the dependency structure. We show that the dependency features can be trained on the dependency graphs, and we show that the best results on the CoNLL-2009 tasks.

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

We present a neural parsing model that generates the input sequence of a target-side constituency trees, and it uses features a single sentence representation. The results show that the performance of our parser can be trained on the basis of the parse tree and a state-of-the-art constituency parser that is a significant improvement over the parser performance.

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

We propose a neural model for neural sequence-tosequence neural models that use neural models to learn continuous representations of a sentence. We show that our model outperforms the state-of-the-art on several language-learning tasks. We empirically introduce a novel neural network model for neural language modeling. Our model significantly outperforms the baseline NNJM on test data, and also significantly outperforms baseline models by using word embeddings.

Composing Finite State Transducers on GPUs
Arturo Argueta and David Chiang

We present a finite-finite finite-state transducer to estimate the flexibility of finite-state and finite state transducers of finite state transducers. We also present a finite-state transducer and finite state transducers, and show that it is possible to train state-of-the-art finite-state transducer.

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

This paper discusses the challenges of training data and tuning for the conversion of the training data. We propose a method for conversion to create a large number of differing data and apply it to the conversion of the training data. We also attempt to reduce the number of training data to reduce the number of target trees, and use the training data

to obtain a small amount of training data. We show that the POS tagger can be trained using a POS tagger and a small number of annotated data that outperform the state of the art on the test data.

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

We propose a novel approach to the task of document understanding and inference in a way of generating coherent stories, events and events, by replacing related events in a single document. We propose a novel probabilistic model for detection, and propose a dynamic programming interface to learn the structure of a directed acyclic graph. We show that the model is able to capture meaningful and pragmatic aspects that perceive the query.

An Empirical Study of Building a Strong Baseline for Constituency Parsing
Jun Suzuki, Sho Takase, Hidetaka Kamigaito, Makoto Morishita, and Masaaki Nagata

We propose an approach to the task of resolving constituency parses, which is a necessary step for natural language processing. We will introduce the first step, using a supervised learning framework to obtain the best possible circumstances of the parser. We show that the proposed method is not scalable enough to parse training data, and we use a parser to distinguish between the dependency graphs and dependency graphs. We conduct experiments on string parsing, and show that the proposed models can be significantly improved by 0.4%, giving high accuracy, surpassing the optimal phrase-based system and the best tagger over the entire parse trees, only on all cases.

Parser Training with Heterogeneous Treebanks
Sara Stymne, Miryam de Lhoneux, Aaron Smith, and Joakim Nivre

We present a new state-of-the-art dependency parser that is trained directly on parse trees. We propose a simple neural parser that can be trained on the small amount of training data. We train the model that can be trained on a small amount of training data. We evaluate the model on the task of parsing, showing that the training parameters with the probability distributions is that the optimized parameters alone. We show that the neural model trained on a small sentence, and leads to significant improvements over the baseline.

Generalized chart constraints for efficient PCFG and TAG parsing
Stefan Grünewald, Sophie Henning, and Alexander Koller

We present a interactive speech translation tool for analyzing the effects of non-speech emotional frequencies. We show that the modification of the word frequency and the effect of the whole corpus is not sufficient to find non-native speakers. We find that this modification is not only on the language, and show that they can be represented as a lexicon as a PCFG surprisal, and to measure the complexity of the corpus as a baseline. We show that this modification is not conditioned on the corpora, and show that they can be used in a corpus-to-Japanese manner, and we see the use of a baseline-based method. We find that a weighted amount of the corpus was actually said to annotate non-finite words, and we show that the use of hand-coded syllables as a stable.

Demo Poster Session 3

Time: *12:30–14:00* Location: Melbourne Room 1 & 2

The SUMMA Platform: A Scalable Infrastructure for Multi-lingual Multi-media Monitoring
Ulrich Germann, Renars Liepins, Guntis Barzdins, Didzis Gosko, Sebastião Miranda, and David Nogueira

This paper aims to investigate multilingual adaptation of multilingual semantic parsers for multilingual language documentation. We introduce a new multilingual corpus for multilingual machine translation, multilingual data sources, and discuss the potential of developing a multilingual parallel corpus for a resource-poor language.

CRUISE: Cold-Start New Skill Development via Iterative Utterance Generation
Yilin Shen, Avik Ray, Abhishek Patel, and Hongxia Jin

We propose a novel approach to automatically identify meanings of sentences in a multi-source framework. We use the model to a word-level learning framework, allowing the model to learn a coherent semantic structure of input sentences that ensure that the learned words are incomplete in the input sentence. We also show how the model is able to generate a complete version of the utterance. We introduce a novel version of the grammars Tree Grammar model and a programming part of a set of utterances that are directed and elementary phrases and the same structure of the target language. We show that the proposed model outperforms the baselines and the state of the art.

Praaline: An Open-Source System for Managing, Annotating, Visualising and Analysing Speech Corpora
George Christodoulides

We present a novel dialogue system for spoken language understanding. We report on a freely available system for annotating speech tags and non-standard transcription. We show that the system is able to generate ASR outputs of a given sentence. We report on our system by a state-of-the-art speech recognition system, and we report on the first holistic error correction that is able to identify the extent to which the agent is likely.

Marian: Fast Neural Machine Translation in C++
Marcin Junczys-Dowmunt, Roman Grundkiewicz, Tomasz Dwojak, Hieu Hoang, Kenneth Heafield, Tom Neckermann, Frank Seide, Ulrich Germann, Alham Fikri Aji, Nikolay Bogoychev, André F. T. Martins, and Alexandra Birch

We present an approach to the task of neural machine translation, and use NMT, and help developers to improve the alignments between the source and target-side translations. We describe the system that can be trained on the ability to create a real-time step, and release a web system to generate large-scale parallel data. We trained on a large-scale parallel corpus, and show that neural models can be trained on par with the best-performing baseline state-of-the-art on the word-based translation task.

DeepPavlov: Open-Source Library for Dialogue Systems
Mikhail Burtsev, Alexander Seliverstov, Rafael Airapetyan, Mikhail Arkhipov, Dilyara Baymurzina, Nickolay Bushkov, Olga Gureenkova, Taras Khakhulin, Yuri Kuratov, Denis Kuznetsov, Alexey Litinsky, Varvara Logacheva, Alexey Lymar, Valentin Malykh, Maxim Petrov, Vadim Polulyakh, Leonid Pugachev, Alexey Sorokin, Maria Vikhreva, and Marat Zaynutdinov

We present open-domain dialogue systems for dialogue systems. We release a dataset of a set of real-world dialogues ranging from real-world dialogue systems, and show that it is beneficial to both automatic and domain adaptation, and a dialogue system for dialogue systems that can be used to predict the reward. We also present an open-source dialogue system to address this new task, and present a unified formalization of this paradigm. Furthermore, the resulting system is able to handle the hard of acceptable responses to reliably learners. We release the handwritten domains and the L1 of the domain-specific domain mismatch between domains and scripts.

RETURNN as a Generic Flexible Neural Toolkit with Application to Translation and Speech Recognition
Albert Zeyer, Tamer Alkhouli, and Hermann Ney

Speech segmentation is a language technology (ASR) is a fundamental task in systems that aim to improve automatic speech recognition systems. We propose a novel approach to efficiently neural machine translation (NMT). We propose a novel approach to ChineseEnglish adaptation neural machine translation (ASR). We propose a novel approach to ChineseEnglish segmentation, but also implements a neural framework for modeling the nested complex language. Experiments show that the model is character-based on a character-based model, and generates translations from compressed derivations, and we also show that our model outperforms state-of-the-art methods.

A Flexible, Efficient and Accurate Framework for Community Question Answering Pipelines
Salvatore Romeo, Giovanni Da San Martino, Alberto Barrón-Cedeño, and Alessandro Moschitti

Community question answering (CQA) is a question-question answering. It is mainly based on the question answering community. However, it is hard to find a question can be paid in the enterprise community. In this paper, we propose a novel approach for question answering (KBQA) that can be used to solve the problem. We propose a novel method for generating named entities in the BioASQ domain, and we propose a novel method for solving the problem of question answering. We compare these two tasks, including both experts and coverage, and various methods of the semantic similarities between questions and factoid question answering questions. We also present a novel question-answering question answering system that incorporates a fine-grained answer selection task. By encoding the question-answer pairs, we show that the proposed method can achieve better results on the task of question answering.

Moon IME: Neural-based Chinese Pinyin Aided Input Method with Customizable Association
Yafang Huang, Zuchao Li, Zhuosheng Zhang, and Hai Zhao

We present a novel method for machine translation and the expansion of a set of pronouns in a language-specific translation. We will describe the first attempt to build a freely available translation system and the use of a parallel corpus. Results show that the proposed approach is easier for the task of combining a more language-specific segmentation. Results show that the proposed approach outperforms the traditional way of developing machine translation and light language-independent tests.

Session 8 Overview – Wednesday, July 18, 2018

	Track A *Semantics 2 (Short)*	**Track B** *Machine Translation, Multilinguality 2 (Short)*	**Track C** *Information Extraction 2 (Short)*	**Track D** *Generation, Summarization (Short)*	**Track E** *Machine Learning, Question Answering (Short)*	**Track F** *Sentiment (Short)*
	Plenary	203–204	210–211	212–213	219	220
14:00	Exploring Semantic Properties of Sentence Embeddings *Zhu, Li, and Melo*	Adaptive Knowledge Sharing in Multi-Task Learning: Improving Low-Resource Neural Machine Translation *Zaremoodi, Buntine, and Haffari*	Enhancing Drug-Drug Interaction Extraction from Texts by Molecular Structure Information *Asada, Miwa, and Sasaki*	Personalized Review Generation By Expanding Phrases and Attending on Aspect-Aware Representations *Ni and McAuley*	Long Short-Term Memory as a Dynamically Computed Element-wise Weighted Sum *Levy, Lee, FitzGerald, and Zettlemoyer*	A Multi-sentiment-resource Enhanced Attention Network for Sentiment Classification *Lei, Yang, Yang, and Liu*
14:15	Scoring Lexical Entailment with a Supervised Directional Similarity Network *Rei, Gerz, and Vulić*	Automatic Estimation of Simultaneous Interpreter Performance *Stewart, Vogler, Hu, Boyd-Graber, and Neubig*	diaNED: Time-Aware Named Entity Disambiguation for Diachronic Corpora *Agarwal, Strötgen, Del Corro, Hoffart, and Weikum*	Learning Simplifications for Specific Target Audiences *Scarton and Specia*	On the Practical Computational Power of Finite Precision RNNs for Language Recognition *Weiss, Goldberg, and Yahav*	Pretraining Sentiment Classifiers with Unlabeled Dialog Data *Shimizu, Shimizu, and Kobayashi*
14:30	Extracting Commonsense Properties from Embeddings with Limited Human Guidance *Yang, Birnbaum, Wang, and Downey*	Polyglot Semantic Role Labeling *Mulcaire, Swayamdipta, and Smith*	Examining Temporality in Document Classification *Huang and Paul*	Split and Rephrase: Better Evaluation and Stronger Baselines *Aharoni and Goldberg*	A Co-Matching Model for Multi-choice Reading Comprehension *Wang, Yu, Jiang, and Chang*	Disambiguating False-Alarm Hashtag Usages in Tweets for Irony Detection *Huang, Chen, and Chen*
14:45	Breaking NLI Systems with Sentences that Require Simple Lexical Inferences *Glockner, Shwartz, and Goldberg*	Learning Cross-lingual Distributed Logical Representations for Semantic Parsing *Zou and Lu*	Personalized Language Model for Query Auto-Completion *Jaech and Ostendorf*	Autoencoder as Assistant Supervisor: Improving Text Representation for Chinese Social Media Text Summarization *Ma, Sun, Lin, and Wang*	Tackling the Story Ending Biases in The Story Cloze Test *Sharma, Allen, Bakhshandeh, and Mostafazadeh*	Cross-Target Stance Classification with Self-Attention Networks *Xu, Paris, Nepal, and Sparks*

Parallel Session 8

Session 8A: Semantics 2 (Short)

Plenary *Chair: Omri Abend*

Exploring Semantic Properties of Sentence Embeddings
Xunjie Zhu, Tingfeng Li, and Gerard de Melo 14:00–14:15

Deep neural network architectures have been widely studied in NLP tasks. We investigate the capability of word embeddings that capture semantic properties of word context and perform context-aware word embeddings. We propose a novel approach to evaluating word embeddings using word embeddings, and introduce a neural network to learn the embeddings of a word. Our model achieves state-of-the-art performance on standard benchmarks.

Scoring Lexical Entailment with a Supervised Directional Similarity Network
Marek Rei, Daniela Gerz, and Ivan Vulić 14:15–14:30

In this paper we present the first work on identifying the semantic similarity of the two sentences. We propose a method for determining the similarity between the two sentences. We evaluate the model on a small set of lexical datasets, and show that these embeddings are applied to the word clusters. Furthermore, we present the first place in the task of predicting the paraphrase pairs of the compound multi-document context.

Extracting Commonsense Properties from Embeddings with Limited Human Guidance
Yiben Yang, Larry Birnbaum, Ji-Ping Wang, and Doug Downey 14:30–14:45

We present a novel probabilistic model that learns the distributed representation of a skip-gram language, and find that it learns to predict the meaning of the current neural network. We find that this method learns significant correlations between supervised models, to approximate a bias reward function with a small number of possible mappings. We show that this model learns temporal properties with justifications, and extends the model to train ASR models to ensure the usual predictive cost. Additionally, we show that it is effective for finding unknown words in a single language. We show that it outperforms strong baselines on two datasets, and show that it can be used to train classifiers for ASR models.

Breaking NLI Systems with Sentences that Require Simple Lexical Inferences
Max Glockner, Vered Shwartz, and Yoav Goldberg 14:45–15:00

We present a novel approach to unsupervised lexical simplification, and develop an approach to automatically correcting human utterances. We empirically evaluate the quality of the different types of automatically induced active judgments, and propose a novel rule-based method to identify relevant sentences that distinguish between them. We also present a novel method to identify the expressions of the morphological tags, and then apply it to explore the flexibility of stacking. We validate the performance of manually annotated datasets, achieving an accuracy of 79%, and outperforms the baselines reported on this task.

Session 8B: Machine Translation, Multilinguality 2 (Short)

203–204 *Chair: Yulia Tsvetkov*

Adaptive Knowledge Sharing in Multi-Task Learning: Improving Low-Resource Neural Machine Translation
Poorya Zaremoodi, Wray Buntine, and Gholamreza Haffari 14:00–14:15

We present a novel neural language model that can be used for neural machine translation. We propose a novel task of multi-document summarization, and we propose a novel way to learn continuous representations of target words. We show that the model is a novel attention-based model that is based on knowledge distillation. We show that our model has a significant improvement over the state-of-the-art baseline.

Automatic Estimation of Simultaneous Interpreter Performance
Craig Stewart, Nikolai Vogler, Junjie Hu, Jordan Boyd-Graber, and Graham Neubig 14:15–14:30

This paper investigates the impact of word frequency and lexical substitution, and report results for the task of word sense disambiguation and classification. We find that the word and the word frequency can be used as a frequency of the top-performing measure. We report the performance of the resulting variants of agreement and about agreement and better results.

Polyglot Semantic Role Labeling
Phoebe Mulcaire, Swabha Swayamdipta, and Noah A. Smith 14:30–14:45

Semantic role labeling (SRL), semantic role labeling, which is a hard component for semantic role labeling (SRL). This paper discusses the role of the semantic role structure in the context of FrameNet frames. We propose a novel approach to FrameNet semantic roles, which focus on FrameNet SRL. We define transformation and coverage and advance semantic role labeling, which we will focus on the role of SRL.

Learning Cross-lingual Distributed Logical Representations for Semantic Parsing
Yanyan Zou and Wei Lu 14:45–15:00

We propose a novel neural semantic parsing model that learns latent semantic representations of semantic parses. We show that this approach learns to learn distributed representations, using a model that performs well-formed logical forms, and then carries the number of semantic information. We show that this framework can efficiently learn distributed representations of sentences into a single vector representation of semantic chains. We show that this model can efficiently learn distributed representations of sentences into a single vector representation of semantic trees, and we report a strong state-of-the-art performance on benchmark datasets. The resulting parsers can be used to learn the mapping of dense semantic representations in a semantic parser.

Session 8C: Information Extraction 2 (Short)

210–211 *Chair: Radu Florian*

Enhancing Drug-Drug Interaction Extraction from Texts by Molecular Structure Information
Masaki Asada, Makoto Miwa, and Yutaka Sasaki 14:00–14:15

This paper investigates the use of a machine learning system that extracts relations learnt by extracting user profiles extracted from a large collection of documents. We propose a method for extracting vital information extraction from documents without the keywords to extract keywords from the text. Our approach is based on a single tree-based indexing algorithm. It is also applied to a collection of structured texts. We evaluate the proposed approach on the TempEval news dataset and manually annotate the three levels of biology. Experimental results show that our approach outperforms the baseline system by up to 15%.

diaNED: Time-Aware Named Entity Disambiguation for Diachronic Corpora
Prabal Agarwal, Jannik Strötgen, Luciano Del Corro, Johannes Hoffart, and Gerhard Weikum 14:15–14:30

We present a novel approach to obtaining correct entity mentions in a unified space of text. We measure the feasibility of developing a deterministic tool for mining the number of steps in a supervised setting. We train a weakly supervised classifier that estimates a label distribution in labels for the set of a set of senses. We evaluate the effects of using a number of parameters, showing that our approach outperforms the state-of-the-art methods, and the usefulness of manual annotation.

Examining Temporality in Document Classification
Xiaolei Huang and Michael J. Paul 14:30–14:45

In this paper, we propose a novel approach to extractive text classification, inspired by the document classification task. The proposed approach is based on a word-level representation of the document representation. The model is evaluated on a corpus of running corpus-specific features, and show that our approach achieves state-of-the-art performance on several benchmark datasets.

Personalized Language Model for Query Auto-Completion
Aaron Jaech and Mari Ostendorf 14:45–15:00

Summarization of social media is a language-specific activity, that can be used to learn a target language. Previous work on language modeling and language modeling can be trained with a model that can be trained on a large scale and that it can be used to tailor the interpretation of word meaning. In this paper, we propose a novel language-independent approach to personalized language-independent inference, develop a novel language-independent approach to personalized language-transformation learning, and incorporates a latent variable-mapping between the source language. We show that the latent variable model is trained on a small set of human user-provided questions.

Session 8D: Generation, Summarization (Short)

212–213 *Chair: Hiroya Takamura*

Personalized Review Generation By Expanding Phrases and Attending on Aspect-Aware Representations

Jianmo Ni and Julian McAuley 14:00–14:15

We propose a novel neural network based on convolutional neural networks that can be used to extract latent attributes of the structure of target text. We propose a novel bidirectional recurrent neural network with Attention (Open Domain), which we use convolutional neural networks to detect the latent variables of the target sentence. We propose a novel neural network that learns to embed both meaning, i.e., phrases and captions that that can be used to adapt to personalized reviews. The experiments show that it is effective, while generating informative knowledge can be effectively learned from the area of domain adaptation.

Learning Simplifications for Specific Target Audiences

Carolina Scarton and Lucia Specia 14:15–14:30

We propose a novel approach to learning a sentence-based semantic representations of sentence pairs. We show that pretraining a simple word-representation method, which can be thought to be conform to the compressed parts, and it is sensitive to the problem of inference. We propose a novel approach to learning coarse-source features, adopting the first two components that we focused on an attentional phrase-learning model. We show that the proposed approach can significantly improve over the baselines.

Split and Rephrase: Better Evaluation and Stronger Baselines

Roee Aharoni and Yoav Goldberg 14:30–14:45

We propose two approaches to evaluate the quality of the crowd-defined word embeddings. We propose a novel entailment method that explores the order of multiple semantic relatedness and the entailment of each word, and then use the information of the word-to-text structure. We show that the hybrid asymmetry model in the context of the context and the model can be trained on the same dataset. We find that the proposed methods outperform previously baselines and a number of baselines.

Autoencoder as Assistant Supervisor: Improving Text Representation for Chinese Social Media Text Summarization

Shuming Ma, Xu Sun, Junyang Lin, and Houfeng Wang 14:45–15:00

Many tasks are a promising approach for language modeling. Existing approaches to extractive text summarization is a challenging problem because the solution is solved to train the traditional approach. However, they suffer from limited challenges, but also retaining the connections of different models. In this paper, we propose a novel approach to learn distributed representations of words to capture the context of the input. We propose a neural network model to learn abstractive summaries and learn the tensor architecture to compute the structure of the original document. We propose a novel LSTM model that allows us to learn a large fraction of informative words. Experiments show that our approach outperforms the previous state-of-the-art on the semi-supervised dataset.

Session 8E: Machine Learning, Question Answering (Short)

219 *Chair: Massimo Piccardi*

Long Short-Term Memory as a Dynamically Computed Element-wise Weighted Sum
Omer Levy, Kenton Lee, Nicholas FitzGerald, and Luke Zettlemoyer 14:00–14:15

In this paper we present an analysis of the model that jointly employs a sequence of a recurrent neural network, a natural language-based grammar-oriented dialogue. It is able to model the semantics of the full holistic head of input. This model updates the memory and the structured semantics, and then learning the next part of the semantics of the dialogue. The model is designed to capture the sequential structure of a clause. It is a pre-processing step, therefore we instead the use of sequential RNNs and its ability to capture sequential information. The model is fed to neural networks and a sequence-to-sequence model, which can be deployed for a given input.

On the Practical Computational Power of Finite Precision RNNs for Language Recognition
Gail Weiss, Yoav Goldberg, and Eran Yahav 14:15–14:30

We present a novel approach for evaluating sentence-based natural language processing (NLP). We show that the proposed model is able to encode the generalizations of the semantic roles of a sentence. We show that the proposed model is able to encode the dependency structure of a sentence. We show that the proposed model is able to capture the grammatical regularities in the training data.

A Co-Matching Model for Multi-choice Reading Comprehension
Shuohang Wang, Mo Yu, Jing Jiang, and Shiyu Chang 14:30–14:45

Modeling information is a critical research problem in natural language understanding. This paper studies the problem of multi-task learning and inference (i.e., a binary), which is a more important task. We propose a novel neural model that extends the structured inference, which is a number of parameters that have the capacity of a document. We show that our model can effectively improve the quality of this dataset. Results show that the proposed models can significantly improve the quality of the generated texts by the use of a reinforcement learning.

Tackling the Story Ending Biases in The Story Cloze Test
Rishi Sharma, James Allen, Omid Bakhshandeh, and Nasrin Mostafazadeh 14:45–15:00

This paper reports on the evaluation of the proposed neural network architecture for building natural language responses. We focus on neural networks, and propose a model for recognizing textual entailment. We focused on neural networks, and propose a model that can be trained on the basis of neural responses. We show that the proposed approach can significantly improve performance.

Session 8F: Sentiment (Short)

220 *Chair: Ion Androutsopoulos*

A Multi-sentiment-resource Enhanced Attention Network for Sentiment Classification
Zeyang Lei, Yujiu Yang, Min Yang, and Yi Liu 14:00–14:15

We propose a novel task of sentiment classification from multilingual document classification that can be used to predict the sentiment polarity of a document. We propose a novel attention mechanism for sentiment classification and model the relationship between the source and target word context, and propose a novel attention-based attention model that can fuse the attention-based attention model that can fuse the attention-level attention. Experiments show that the proposed model outperforms the state-of-the-art on two tasks.

Pretraining Sentiment Classifiers with Unlabeled Dialog Data
Toru Shimizu, Nobuyuki Shimizu, and Hayato Kobayashi 14:15–14:30

In this paper, we propose a novel approach to rapidly use sentiment analysis of training data (e.g. regression). We propose a novel approach to automatically learn the linguistic features of opinions and their interactions. We show that the proposed models can be successfully trained to predict the whole study of the samples. Results show that the proposed models can be successfully trained to predict the sentiment of a target tweet.

Disambiguating False-Alarm Hashtag Usages in Tweets for Irony Detection
Hen-Hsen Huang, Chiao-Chen Chen, and Hsin-Hsi Chen 14:30–14:45

The detection of verbal multiword expressions (ED) in social media has led to significant progress in detection. We investigate the task of detecting sarcasm framing in an social media essays. We propose a novel approach to detecting metaphors, which encode informative and identifying communicative messages. We show that temporal annotations of tweets can be trained effectively using a single tweet classifier and trending classification tasks. We find that the use of hashtags can be trained effectively for a hate speech detection.

Cross-Target Stance Classification with Self-Attention Networks
Chang Xu, Cecile Paris, Surya Nepal, and Ross Sparks 14:45–15:00

Existing studies on stance classification in text classification and stance classification are widely used in natural language processing (NLP) tasks. In this paper, we propose a novel approach for stance classification using a succinct model, which is a key step towards understanding stance in text classification. We collect a new dataset of approximately political clusters, which we introduce a new dataset of multi-document text classification. We show that this is encoded in a document classification task, significantly outperforming state-of-the-art baselines.

Session 9: Best Paper Session (Sponsored by Apple & Amazon)

The process for selecting best papers and honourable mentions

The Program Committee Co-Chairs (PCs) have defined a multi-step process. Area Chairs (ACs) were asked to select a number of top papers in their areas satisfying as many as possible of the following criteria:

- high quality
- nominated for the award by at least one primary reviewer
- bringing disruptive ground-breaking innovation as compared to the current mainstream

ACs re-read their finalists and discussed among themselves the merits of the nominee's work with the help of the primary reviews. ACs then submitted the papers to the PCs along with their selection decisions. PCs balanced ACs' nominations for diversity and representativeness among areas and the review consistency. They prepared the papers in Softconf for best-paper reviewing and selection. There were 52 best paper candidates.

In parallel, PCs formed the best paper selection committee (BPC) from 22 experts in the field with a mix of expertise and backgrounds and at a good seniority level. In case of COIs, the BPC member was excluded from the further evaluation process. BPC members reviewed 6-8 papers each and provided a short review with respect to the best paper criteria.

Based on BPC recommendations, there were about 20 papers left in the pool. PCs then re-read those papers and discussed their particular merits. Finally, 6 long papers and 2 short papers were selected as honourable mentions. For the best papers, 3 long papers and 2 short papers were selected for presentation in the closing conference session.

The selected honourable mentions and best papers emphasize the diversity of the ACL in terms of research questions, methods, and interdisciplinarity.

Best Long Papers

- *Finding syntax in human encephalography with beam search.* John Hale, Chris Dyer, Adhiguna Kuncoro and Jonathan Brennan.
- *Learning to Ask Good Questions: Ranking Clarification Questions using Neural Expected Value of Perfect Information.* Sudha Rao and Hal Daumé III.
- *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.

Best Short Papers

- *Know What You Don't Know: Unanswerable Questions for SQuAD.* Pranav Rajpurkar, Robin Jia and Percy Liang.
- *'Lighter' Can Still Be Dark: Modeling Comparative Color Descriptions.* Olivia Winn and Smaranda Muresan.

Session 9: Best Paper Session

Plenary *Chair: Iryna Gurevych, Yusuke Miyao*

Know What You Don't Know: Unanswerable Questions for SQuAD
Pranav Rajpurkar, Robin Jia, and Percy Liang 15:30–15:45

Question answering (QA) is a common understanding of the task. Recently, reinforcement learning and question learning can effectively extract visual information from the text. However, it is difficult to find pieces of such questions that can be trained on the data to balance the expertise to the question. We propose a novel approach to automatically learn textual information from the context of a dialogue system. We show that our model outperforms the state-of-art automated automated QA systems, and outperforms flat models.

'Lighter' Can Still Be Dark: Modeling Comparative Color Descriptions
Olivia Winn and Smaranda Muresan 15:45–16:00

We present a novel approach for investigating the effect of predicting the ontological status of a word (MWE). We model adjective verbs, such as a noun, roughly, red, and analyze the use of a word. We propose a new set of words that influence the meaning of the meaning of the verb is less likely to be perfect. We propose a new set of words that influence the meaning that the meanings of the verb is more likely to be perfect. We show that the resulting embeddings can be successfully applied to a wide range of semantic spaces.

Finding syntax in human encephalography with beam search
John Hale, Chris Dyer, Adhiguna Kuncoro, and Jonathan Brennan 16:00–16:25

We present a interactive speech translation tool for analyzing the effects of non-speech emotional frequencies. We show that the modification of the word frequency and the effect of the whole corpus is not sufficient to find non-native speakers. We find that this modification is not only on the language, and show that they can be represented as a lexicon as a PCFG surprisal, and to measure the complexity of the corpus as a baseline. We show that this modification is not conditioned on the corpora, and show that they can be used in a corpus-to-Japanese manner, and we see the use of a baseline-based method. We find that a weighted amount of the corpus was actually said to annotate non-finite words, and we show that the use of hand-coded syllables as a stable.

Learning to Ask Good Questions: Ranking Clarification Questions using Neural Expected Value of Perfect Information
Sudha Rao and Hal Daumé III 16:25–16:50

We propose a novel approach to extractive text understanding. We focus on the problem of automatically learned tuples that there are a set of structured data. We propose a novel approach to automatically learn latent operations and reinforcement learning algorithms. We show that both generative models can learn context-sensitive and differentiable predictive models. We show that the proposed models can incorporate task-specific and scalable learning models. We show that the proposed models can incorporate knowledge-rich and generalizable predictive models.

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 16:50–17:15

We present a novel approach that draws on a large-scale pseudo-textual data. We propose a novel variant of compositionality in a supervised setting, which is especially useful for this task, and it is scalable, training data, and it is scalable, training data, and a novel classifier that uses a logistic regression classifier that is trained on an annotated dataset. We trained on a large scale parallel corpus and show that it outperforms the state-of-the-art on the Europarl clinical abbreviation.

6 Workshops

Thursday

Friday

BioNLP: BioNLP 2018 Workshop

Organizers: *Dina Demner-Fushman, Kevin Bretonnel Cohen, Sophia Ananiadou, and Junichi Tsujii*
Venue: 207

Thursday, July 19, 2018

09:00–09:15 **Opening remarks**

09:15–10:30 **Session 1: Clinical NLP**

09:15–09:30 Embedding Transfer for Low-Resource Medical Named Entity Recognition: A Case Study on Patient Mobility
Denis Newman-Griffis and Ayah Zirikly

09:30–09:45 Multi-task learning for interpretable cause of death classification using key phrase prediction
Serena Jeblee, Mireille Gomes, and Graeme Hirst

09:45–10:00 Identifying Risk Factors For Heart Disease in Electronic Medical Records: A Deep Learning Approach
Thanat Chokwijitkul, Anthony Nguyen, Hamed Hassanzadeh, and Siegfried Perez

10:00–10:15 Keyphrases Extraction from User-Generated Contents in Healthcare Domain Using Long Short-Term Memory Networks
Ilham Fathy Saputra, Rahmad Mahendra, and Alfan Farizki Wicaksono

10:15–10:30 Identifying Key Sentences for Precision Oncology Using Semi-Supervised Learning
Jurica Ševa, Martin Wackerbauer, and Ulf Leser

10:30–11:00 **Coffee Break**

11:00–12:00 **Session 2: Foundations**

11:00–11:15 Ontology alignment in the biomedical domain using entity definitions and context
Lucy Wang, Chandra Bhagavatula, Mark Neumann, Kyle Lo, Chris Wilhelm, and Waleed Ammar

11:15–11:30 Sub-word information in pre-trained biomedical word representations: evaluation and hyper-parameter optimization
Dieter Galea, Ivan Laponogov, and Kirill Veselkov

11:30–11:45 PICO Element Detection in Medical Text via Long Short-Term Memory Neural Networks
Di Jin and Peter Szolovits

11:45–12:00 Coding Structures and Actions with the COSTA Scheme in Medical Conversations
Nan Wang, Yan Song, and Fei Xia

12:00–13:30 **Lunch break**

13:30–14:30 **Invited Talk: "Automating systematic reviews: progress and challenges" – Paul Glasziou**

14:30–15:30 **Session 3 Literature mining and retrieval; Question Answering**

14:30–14:45 A Neural Autoencoder Approach for Document Ranking and Query Refinement in Pharmacogenomic Information Retrieval
Jonas Pfeiffer, Samuel Broscheit, Rainer Gemulla, and Mathias Göschl

14:45–15:00 Biomedical Event Extraction Using Convolutional Neural Networks and Dependency Parsing
Jari Björne and Tapio Salakoski

15:00–15:15 BioAMA: Towards an End to End BioMedical Question Answering System
Vasu Sharma, Nitish Kulkarni, Srividya Pranavi, Gabriel Bayomi, Eric Nyberg, and Teruko Mitamura

15:15–15:30 Phrase2VecGLM: Neural generalized language model–based semantic tagging for complex query reformulation in medical IR
Manirupa Das, Eric Fosler-Lussier, Simon Lin, Soheil Moosavinasab, David Chen, Steve Rust, Yungui Huang, and Rajiv Ramnath

15:30–16:00 **Coffee Break**

16:00–16:15 **Invited Presentation: "A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature"–Ben Nye**

16:15–18:00 **Poster Session**

- Convolutional neural networks for chemical-disease relation extraction are improved with character-based word embeddings
Dat Quoc Nguyen and Karin Verspoor
- Domain Adaptation for Disease Phrase Matching with Adversarial Networks
Miaofeng Liu, Jialong Han, Haisong Zhang, and Yan Song
- Predicting Discharge Disposition Using Patient Complaint Notes in Electronic Medical Records
Mohamad Salimi and Alla Rozovskaya
- Bacteria and Biotope Entity Recognition Using A Dictionary-Enhanced Neural Network Model
Qiuyue Wang and Xiaofeng Meng
- SingleCite: Towards an improved Single Citation Search in PubMed
Lana Yeganova, Donald C Comeau, Won Kim, W John Wilbur, and Zhiyong Lu
- A Framework for Developing and Evaluating Word Embeddings of Drug-named Entity
Mengnan Zhao, Aaron J. Masino, and Christopher C. Yang
- MeSH-based dataset for measuring the relevance of text retrieval
Won Gyu Kim, Lana Yeganova, Donald C Comeau, W John Wilbur, and Zhiyong Lu
- CRF-LSTM Text Mining Method Unveiling the Pharmacological Mechanism of Off-target Side Effect of Anti-Multiple Myeloma Drugs
Kaiyin Zhou, Sheng Zhang, Xiangyu Meng, Qi Luo, Yuxing Wang, Ke Ding, Yukun Feng, Mo Chen, Kevin Bretonnel Cohen, and Jingbo Xia
- Prediction Models for Risk of Type-2 Diabetes Using Health Claims
Masatoshi Nagata, Kohichi Takai, Keiji Yasuda, Panikos Heracleous, and Akio Yoneyama
- On Learning Better Embeddings from Chinese Clinical Records: Study on Combining In-Domain and Out-Domain Data
Yaqiang Wang, Yunhui Chen, Hongping Shu, and Yongguang Jiang
- Investigating Domain-Specific Information for Neural Coreference Resolution on Biomedical Texts
Long Trieu, Nhung Nguyen, Makoto Miwa, and Sophia Ananiadou
- Toward Cross-Domain Engagement Analysis in Medical Notes
Sara Rosenthal and Adam Faulkner

CogCL: The Eighth Workshop on Cognitive Aspects of Computational Language Learning and Processing

Organizers: *Marco Idiart, Alessandro Lenci, Thierry Poibeau, and Aline Villavicencio*

Venue: 218

Thursday, July 19, 2018

09:00–09:10 **Welcome and Opening Session**

09:10–09:30 **Session I - Semantics**

09:10–09:30 Predicting Brain Activation with WordNet Embeddings
João António Rodrigues, Ruben Branco, João Silva, Chakaveh Saedi, and António Branco

09:30–10:30 **Invited Talk I**

10:30–11:00 **Coffee Break**

11:00–11:50 **Session II - Production**

11:00–11:20 Do Speakers Produce Discourse Connectives Rationally?
Frances Yung and Vera Demberg

11:20–11:50 Language Production Dynamics with Recurrent Neural Networks
Jesús Calvillo and Matthew Crocker

11:50–12:30 **Poster Session**

11:50–12:30 Multi-glance Reading Model for Text Understanding
Pengcheng Zhu, Yujiu Yang, Wenqiang Gao, and Yi Liu

11:50–12:30 Predicting Japanese Word Order in Double Object Constructions
Masayuki Asahara, Satoshi Nambu, and Shin-Ichiro Sano

11:50–12:30 Affordances in Grounded Language Learning
Stephen McGregor and KyungTae Lim

12:30–14:00 **Lunch**

14:00–15:00 **Invited Talk II**

15:00–15:30 **Session III - Processing**

15:00–15:30 Rating Distributions and Bayesian Inference: Enhancing Cognitive Models of Spatial Language Use
Thomas Kluth and Holger Schultheis

15:30–16:00 **Coffee Break**

16:00–17:00 **Session IV - Syntax and Parsing**

16:00–16:30 The Role of Syntax During Pronoun Resolution: Evidence from fMRI
Jixing Li, Murielle Fabre, Wen-Ming Luh, and John Hale

16:30–17:00 A Sound and Complete Left-Corner Parsing for Minimalist Grammars
Miloš Stanojević and Edward Stabler

17:00–17:30 **Panel, Business Meeting and Closing Session**

DeepLo: Workshop on Deep Learning Approaches for Low Resource NLP

Organizers: *Reza Haffari, Colin Cherry, George Foster, Shahram Khadivi, and Bahar Salehi*
Venue: 211

Thursday, July 19, 2018

09:00–09:10 **Opening Remarks**

09:10–09:50 **Invited Talk (Stefan Riezler)**

09:50–10:30 **Invited Talk (Sujith Ravi)**

10:30–11:00 **Coffee Break**

11:00–12:40 **Oral Presentations**

11:00–11:25 Phrase-Based & Neural Unsupervised Machine Translation
Guillaume Lample, Myle Ott, Alexis Conneau, Ludovic Denoyer, and Marc'Aurelio Ranzato

11:25–11:50 Character-level Supervision for Low-resource POS Tagging
Katharina Kann, Johannes Bjerva, Isabelle Augenstein, Barbara Plank, and Anders Søgaard

11:50–12:15 Training a Neural Network in a Low-Resource Setting on Automatically Annotated Noisy Data
Michael A. Hedderich and Dietrich Klakow

12:15–12:40 Exploiting Cross-Lingual Subword Similarities in Low-Resource Document Classification
Mozhi Zhang, Yoshinari Fujinuma, and Jordan Boyd-Graber

12:40–14:00 **Lunch Break**

14:00–15:30 **Poster Session**

- Multi-task learning for historical text normalization: Size matters
Marcel Bollmann, Anders Søgaard, and Joachim Bingel
- Compositional Language Modeling for Icon-Based Augmentative and Alternative Communication
Shiran Dudy and Steven Bedrick
- Multimodal Neural Machine Translation for Low-resource Language Pairs using Synthetic Data
Koel Dutta Chowdhury, Mohammed Hasanuzzaman, and Qun Liu
- Morphological neighbors beat word2vec on the long tail
Clayton Greenberg, Mittul Singh, and Dietrich Klakow
- Multi-Task Active Learning for Neural Semantic Role Labeling on Low Resource Conversational Corpus
Fariz Ikhwantri, Samuel Louvan, Kemal Kurniawan, Bagas Abisena, Valdi Rachman, Alfan Farizki Wicaksono, and Rahmad Mahendra
- Domain Adapted Word Embeddings for Improved Sentiment Classification
Prathusha K Sarma, Yingyu Liang, and Bill Sethares
- Investigating Effective Parameters for Fine-tuning of Word Embeddings Using Only a Small Corpus
Kanako Komiya and Hiroyuki Shinnou

- Dependency Parsing of Code-Switching Data with Cross-Lingual Feature Representations
  KyungTae Lim, Niko Partanen, Michael Rießler, and Thierry Poibeau
- Semi-Supervised Learning with Auxiliary Evaluation Component for Large Scale e-Commerce Text Classification
  Mingkuan Liu, Musen Wen, Selcuk Kopru, Xianjing Liu, and Alan Lu
- Low-rank passthrough neural networks
  Antonio Valerio Miceli Barone
- Embedding Transfer for Low-Resource Medical Named Entity Recognition: A Case Study on Patient Mobility
  Denis Newman-Griffis and Ayah Zirikly

15:30–16:00 **Coffee Break**

16:00–16:40 **Invited Talk (Trevor Cohn)**

16:40–17:40 **Panel Discussion**

17:40–17:55 **Closing Remarks**

MSR: The First Workshop on Multilingual Surface Realization

Organizers: *Simon Mille, Anja Belz, Bernd Bohnet, Emily Pitler, and Leo Wanner*
Venue: 216

Thursday, July 19, 2018

08:45–09:00 **Opening Remarks**

09:00–10:00 **Invited Talk (Hadar Shemtov)**

10:00–10:30 **Shared Task overview and results**

- The First Multilingual Surface Realisation Shared Task (SR'18): Overview and Evaluation Results
  Simon Mille, Anja Belz, Bernd Bohnet, Yvette Graham, Emily Pitler, and Leo Wanner

10:30–11:00 **Coffee Break**

11:00–12:30 **Oral Presentations**

- BinLin: A Simple Method of Dependency Tree Linearization
  Yevgeniy Puzikov and Iryna Gurevych
- IIT (BHU) Varanasi at MSR-SRST 2018: A Language Model Based Approach for Natural Language Generation
  Shreyansh Singh, Ayush Sharma, Avi Chawla, and A.K. Singh
- Surface Realization Shared Task 2018 (SR18): The Tilburg University Approach
  Thiago Castro Ferreira, Sander Wubben, and Emiel Krahmer

12:30–13:45 **Lunch**

13:45–14:15 **Oral Presentation**

- The OSU Realizer for SRST '18: Neural Sequence-to-Sequence Inflection and Incremental Locality-Based Linearization
  David King and Michael White

14:15–15:30 **Poster Session**

- AX Semantics' Submission to the Surface Realization Shared Task 2018
  Andreas Madsack, Johanna Heininger, Nyamsuren Davaasambuu, Vitaliia Voronik, Michael Käufl, and Robert Weißgraeber
- NILC-SWORNEMO at the Surface Realization Shared Task: Exploring Syntax-Based Word Ordering using Neural Models
  Marco Antonio Sobrevilla Cabezudo and Thiago Pardo
- The DipInfo-UniTo system for SRST 2018
  Valerio Basile and Alessandro Mazzei
- Generating High-Quality Surface Realizations Using Data Augmentation and Factored Sequence Models
  Henry Elder and Chris Hokamp

15:30–16:00 **Coffee Break**

16:00–17:30 **Panel/Discussions**

NLPTEA: The 5th Workshop on Natural Language Processing Techniques for Educational Applications

Organizers: *Yuen-Hsien Tseng, Hsin-Hsi Chen, Vincent Ng, and Mamoru Komachi*

Venue: 217

Thursday, July 19, 2018

09:20–09:30 **Opening Remarks**

09:30–10:30 **Invited Talk**

09:30–10:30 **Multi-word Expressions in Second Language Learning (Yuji Matsumoto)**

10:30–11:00 **Coffee Break**

11:00–12:40 **Regular Paper Session**

11:00–11:20 Generating Questions for Reading Comprehension using Coherence Relations
Takshak Desai, Parag Dakle, and Dan Moldovan

11:20–11:40 Syntactic and Lexical Approaches to Reading Comprehension
Henry Lin

11:40–12:00 Feature Optimization for Predicting Readability of Arabic L1 and L2
Hind Saddiki, Nizar Habash, Violetta Cavalli-Sforza, and Muhamed Al Khalil

12:00–12:20 A Tutorial Markov Analysis of Effective Human Tutorial Sessions
Nabin Maharjan and Vasile Rus

12:20–12:40 Thank "Goodness"! A Way to Measure Style in Student Essays
Sandeep Mathias and Pushpak Bhattacharyya

12:40–14:10 **Lunch**

14:10–15:30 **Shared Task Session**

14:10–14:30 Overview of NLPTEA-2018 Share Task Chinese Grammatical Error Diagnosis
Gaoqi Rao, Qi Gong, Baolin Zhang, and Endong Xun

14:30–14:45 Chinese Grammatical Error Diagnosis using Statistical and Prior Knowledge driven Features with Probabilistic Ensemble Enhancement
Ruiji Fu, Zhengqi Pei, Jiefu Gong, Wei Song, Dechuan Teng, Wanxiang Che, Shijin Wang, Guoping Hu, and Ting Liu

14:45–15:00 A Hybrid System for Chinese Grammatical Error Diagnosis and Correction
Chen Li, Junpei Zhou, Zuyi Bao, Hengyou Liu, Guangwei Xu, and Linlin Li

15:00–15:15 LingCASS Solution to the NLP-TEA CGED Shared Task 2018
Qinan Hu, Yongwei Zhang, Fang Liu, and Yueguo Gu

15:15–15:30 Chinese Grammatical Error Diagnosis Based on Policy Gradient LSTM Model
Changliang Li and Ji Qi

15:30–16:00 **Coffee Break**

16:00–17:00 **Poster Session**

- The Importance of Recommender and Feedback Features in a Pronunciation Learning Aid
Dzikri Fudholi and Hanna Suominen
- Selecting NLP Techniques to Evaluate Learning Design Objectives in Collaborative Multi-perspective Elaboration Activities
Aneesha Bakharia

- Augmenting Textual Qualitative Features in Deep Convolution Recurrent Neural Network for Automatic Essay Scoring
  Tirthankar Dasgupta, Abir Naskar, Lipika Dey, and Rupsa Saha
- Joint learning of frequency and word embeddings for multilingual readability assessment
  Dieu-Thu Le, Cam-Tu Nguyen, and Xiaoliang Wang
- MULLE: A grammar-based Latin language learning tool to supplement the classroom setting
  Herbert Lange and Peter Ljunglöf
- Textual Features Indicative of Writing Proficiency in Elementary School Spanish Documents
  Gemma Bel-Enguix, Diana Dueñas Chavez, and Arturo Curiel Díaz
- Assessment of an Index for Measuring Pronunciation Difficulty
  Katsunori Kotani and Takehiko Yoshimi
- A Short Answer Grading System in Chinese by Support Vector Approach
  Shih-Hung Wu and Wen-Feng Shih
- From Fidelity to Fluency: Natural Language Processing for Translator Training
  Oi Yee Kwong
- Countering Position Bias in Instructor Interventions in MOOC Discussion Forums
  Muthu Kumar Chandrasekaran and Min-Yen Kan
- Measuring Beginner Friendliness of Japanese Web Pages explaining Academic Concepts by Integrating Neural Image Feature and Text Features
  Hayato Shiokawa, Kota Kawaguchi, Bingcai Han, Takehito Utsuro, Yasuhide Kawada, Masaharu Yoshioka, and Noriko Kando
- Learning to Automatically Generate Fill-In-The-Blank Quizzes
  Edison Marrese-Taylor, Ai Nakajima, Yutaka Matsuo, and Ono Yuichi
- Multilingual Short Text Responses Clustering for Mobile Educational Activities: a Preliminary Exploration
  Yuen-Hsien Tseng, Lung-Hao Lee, Yu-Ta Chien, Chun-Yen Chang, and Tsung-Yen Li
- Chinese Grammatical Error Diagnosis Based on CRF and LSTM-CRF model
  Yujie Zhou, Yinan Shao, and Yong Zhou
- Contextualized Character Representation for Chinese Grammatical Error Diagnosis
  Jianbo Zhao, Si Li, and Zhiqing Lin
- CMMC-BDRC Solution to the NLP-TEA-2018 Chinese Grammatical Error Diagnosis Task
  Zhang Yongwei, Hu Qinan, Liu Fang, and Gu Yueguo
- Detecting Simultaneously Chinese Grammar Errors Based on a BiLSTM-CRF Model
  Yajun Liu, Hongying Zan, Mengjie Zhong, and Hongchao Ma
- A Hybrid Approach Combining Statistical Knowledge with Conditional Random Fields for Chinese Grammatical Error Detection
  Yiyi Wang and Chilin Shih
- CYUT-III Team Chinese Grammatical Error Diagnosis System Report in NLPTEA-2018 CGED Shared Task
  Shih-Hung Wu, JUN-WEI Wang, Liang-Pu Chen, and Ping-Che Yang
- Detecting Grammatical Errors in the NTOU CGED System by Identifying Frequent Subsentences
  Chuan-Jie Lin and Shao-Heng Chen

17:00–17:10 **Closing Remarks**

CALCS: The Third Workshop on Computational Approaches to Linguistic Code-Switching

Organizers: *Gustavo Aguilar, Fahad AlGhamdi, Victor Soto, Thamar Solorio, Mona Diab, and Julia Hirschberg*

Venue: 209

Thursday, July 19, 2018

09:00–10:30 **Session 1 Invited Talk and Oral Presentations**

09:00–09:05 **Opening Remarks (Thamar Solorio)**

09:05–09:50 **Invited Talk: Learning to Codeswitch (Pascale Fung)**

09:50–10:10 Joint Part-of-Speech and Language ID Tagging for Code-Switched Data
Victor Soto and Julia Hirschberg

10:10–10:30 Phone Merging For Code-Switched Speech Recognition
Sunit Sivasankaran, Brij Mohan Lal Srivastava, Sunayana Sitaram, Kalika Bali, and Monojit Choudhury

10:30–11:00 **Coffee Break**

11:00–12:00 **Session 2 Oral Presentations**

11:00–11:20 Improving Neural Network Performance by Injecting Background Knowledge: Detecting Code-switching and Borrowing in Algerian texts
Wafia Adouane, Jean-Philippe Bernardy, and Simon Dobnik

11:20–11:40 Code-Mixed Question Answering Challenge: Crowd-sourcing Data and Techniques
Khyathi Chandu, Ekaterina Loginova, Vishal Gupta, Josef van Genabith, Günter Neuman, Manoj Chinnakotla, Eric Nyberg, and Alan W. Black

11:40–12:00 Transliteration Better than Translation? Answering Code-mixed Questions over a Knowledge Base
Vishal Gupta, Manoj Chinnakotla, and Manish Shrivastava

12:00–13:30 **Lunch Break**

13:30–14:15 **Session 3 Invited Talk**

13:30–14:15 **Invited Talk: Variation in Codeswitched Language: a Psycholinguistic Approach to What, When, and Why (Melinda Fricke)**

14:15–15:30 **Session 4 Poster Session**

- Language Identification and Analysis of Code-Switched Social Media Text
Deepthi Mave, Suraj Maharjan, and Thamar Solorio
- Code-Switching Language Modeling using Syntax-Aware Multi-Task Learning
Genta Indra Winata, Andrea Madotto, Chien-Sheng Wu, and Pascale Fung
- Predicting the presence of a Matrix Language in code-switching
Barbara Bullock, Wally Guzman, Jacqueline Serigos, Vivek Sharath, and Almeida Jacqueline Toribio
- Automatic Detection of Code-switching Style from Acoustics
SaiKrishna Rallabandi, Sunayana Sitaram, and Alan W. Black
- Accommodation of Conversational Code-Choice
Anshul Bawa, Monojit Choudhury, and Kalika Bali

- Language Informed Modeling of Code-Switched Text
  Khyathi Chandu, Thomas Manzini, Sumeet Singh, and Alan W. Black
- GHHT at CALCS 2018: Named Entity Recognition for Dialectal Arabic Using Neural Networks
  Mohammed Attia, Younes Samih, and Wolfgang Maier
- Simple Features for Strong Performance on Named Entity Recognition in Code-Switched Twitter Data
  Devanshu Jain, Maria Kustikova, Mayank Darbari, Rishabh Gupta, and Stephen Mayhew
- Bilingual Character Representation for Efficiently Addressing Out-of-Vocabulary Words in Code-Switching Named Entity Recognition
  Genta Indra Winata, Chien-Sheng Wu, Andrea Madotto, and Pascale Fung
- Named Entity Recognition on Code-Switched Data Using Conditional Random Fields
  Utpal Kumar Sikdar, Biswanath Barik, and Björn Gambäck
- The University of Texas System Submission for the Code-Switching Workshop Shared Task 2018
  Florian Janke, Tongrui Li, Eric Rincón, Gualberto Guzmán, Barbara Bullock, and Almeida Jacqueline Toribio
- Tackling Code-Switched NER: Participation of CMU
  Parvathy Geetha, Khyathi Chandu, and Alan W. Black
- Multilingual Named Entity Recognition on Spanish-English Code-switched Tweets using Support Vector Machines
  Daniel Claeser, Samantha Kent, and Dennis Felske

15:30–16:00 **Coffee Break**

16:00–17:00 **Session 5 Shared Task Talks**

16:00–16:10 Named Entity Recognition on Code-Switched Data: Overview of the CALCS 2018 Shared Task
Gustavo Aguilar, Fahad AlGhamdi, Victor Soto, Mona Diab, Julia Hirschberg, and Thamar Solorio

16:10–16:30 IIT (BHU) Submission for the ACL Shared Task on Named Entity Recognition on Code-switched Data
Shashwat Trivedi, Harsh Rangwani, and Anil Kumar Singh

16:30–16:50 Code-Switched Named Entity Recognition with Embedding Attention
Changhan Wang, Kyunghyun Cho, and Douwe Kiela

16:50–17:00 **Closing Remarks (Victor Soto)**

MRQA: Workshop on Machine Reading for Question Answering

Organizers: *Eunsol Choi, Minjoon Seo, Danqi Chen, Robin Jia, and Jonathan Berant*
Venue: 210

Thursday, July 19, 2018

08:45–09:00 **Opening Remarks**

Session 1

09:00–09:35 **Invited Talk: Phil Blunsom, University of Oxford / Deepmind**

09:35–10:10 **Invited Talk: Sebastian Riedel, University College London**

10:10–10:30 **Best Paper Talk: A Systematic Classification of Knowledge, Reasoning, and Context within the ARC Dataset**

10:30–11:00 **Morning coffee break**

Session 2

11:00–11:35 **Invited Talk: Richard Socher, Salesforce Research**

11:35–12:10 **Invited Talk: Jianfeng Gao, Microsoft Research**

12:10–13:45 **Lunch**

Session 3

13:45–14:20 **Invited Talk: Antoine Bordes, Facebook AI Research**

14:20–15:30 **Poster Session (with a spotlight presentation)**

14:20–15:30 Ruminating Reader: Reasoning with Gated Multi-hop Attention
Yichen Gong and Samuel Bowman

14:20–15:30 Systematic Error Analysis of the Stanford Question Answering Dataset
Marc-Antoine Rondeau and T. J. Hazen

14:20–15:30 A Multi-Stage Memory Augmented Neural Network for Machine Reading Comprehension
Seunghak Yu, Sathish Reddy Indurthi, Seohyun Back, and Haejun Lee

14:20–15:30 Tackling Adversarial Examples in QA via Answer Sentence Selection
Yuanhang Ren, Ye Du, and Di Wang

14:20–15:30 DuReader: a Chinese Machine Reading Comprehension Dataset from Real-world Applications
Wei He, Kai Liu, Jing Liu, Yajuan Lyu, Shiqi Zhao, Xinyan Xiao, Yuan Liu, Yizhong Wang, Hua Wu, Qiaoqiao She, Xuan Liu, Tian Wu, and Haifeng Wang

14:20–15:30 Robust and Scalable Differentiable Neural Computer for Question Answering
Jörg Franke, Jan Niehues, and Alexander Waibel

14:20–15:30 A Systematic Classification of Knowledge, Reasoning, and Context within the ARC Dataset
Michael Boratko, Harshit Padigela, Divyendra Mikkilineni, Pritish Yuvraj, Rajarshi Das, Andrew McCallum, Maria Chang, Achille Fokoue-Nkoutche, Pavan Kapanipathi, Nicholas Mattei, Ryan Musa, Kartik Talamadupula, and Michael Witbrock

14:20–15:30 RECIPE: Applying Open Domain Question Answering to Privacy Policies
Yan Shvartzshanider, Ananth Balashankar, Thomas Wies, and Lakshminarayanan Subramanian

14:20–15:30 Neural Models for Key Phrase Extraction and Question Generation
Sandeep Subramanian, Tong Wang, Xingdi Yuan, Saizheng Zhang, Adam Trischler, and Yoshua Bengio

14:20–15:30 Comparative Analysis of Neural QA models on SQuAD
Soumya Wadhwa, Khyathi Chandu, and Eric Nyberg

14:20–15:30 Adaptations of ROUGE and BLEU to Better Evaluate Machine Reading Comprehension Task
An Yang, Kai Liu, Jing Liu, Yajuan Lyu, and Sujian Li

15:30–16:00 **Afternoon coffee break**

16:00–17:00 **Panel Discussion**

RELNLP: Workshop on Relevance of Linguistic Structure in Neural Architectures for NLP

Organizers: *Georgiana Dinu, Miguel Ballesteros, Avirup Sil, Sam Bowman, Wael Hamza, Anders Søgaard, Tahira Naseem, and Yoav Goldberg*

Venue: 208

Thursday, June 19, 2018

08:50–09:00 **Opening Remarks**

Session 1

09:00–10:00 **Invited Talk: Jason Eisner**

10:00–10:20 Compositional Morpheme Embeddings with Affixes as Functions and Stems as Arguments
Daniel Edmiston and Karl Stratos

10:20–11:00 **Break**

Session 2

11:00–12:00 **Invited Talk: Mark Johnson**

12:00–12:20 Unsupervised Source Hierarchies for Low-Resource Neural Machine Translation
Anna Currey and Kenneth Heafield

12:20–13:30 **Lunch**

13:30–14:30 **Session 3 (Poster)**

- Latent Tree Learning with Differentiable Parsers: Shift-Reduce Parsing and Chart Parsing
  Jean Maillard and Stephen Clark
- Syntax Helps ELMo Understand Semantics: Is Syntax Still Relevant in a Deep Neural Architecture for SRL?
  Emma Strubell and Andrew McCallum
- Subcharacter Information in Japanese Embeddings: When Is It Worth It?
  Marzena Karpinska, Bofang Li, Anna Rogers, and Aleksandr Drozd
- A neural parser as a direct classifier for head-final languages
  Hiroshi Kanayama, Masayasu Muraoka, and Ryosuke Kohita
- Syntactic Dependency Representations in Neural Relation Classification
  Farhad Nooralahzadeh and Lilja Øvrelid

14:30–15:30 **Invited Talk: Emily Bender**

15:30–16:00 **Break**

Session 4

16:00–17:00 **Invited Talk: Chris Dyer**

17:00–18:00 **Panel discussion**

ECONLP: The First Workshop on Economics and Natural Language Processing

Organizers: *Udo Hahn, Véronique Hoste, and Ming-Feng Tsai*
Venue: 207

Friday, July 20, 2018

09:00–09:30 **Introduction to the ECONLP Workshop (Udo Hahn)**

09:30–10:00 Economic Event Detection in Company-Specific News Text
Gilles Jacobs, Els Lefever, and Véronique Hoste

10:00–10:30 Causality Analysis of Twitter Sentiments and Stock Market Returns
Narges Tabari, Piyusha Biswas, Bhanu Praneeth, Armin Seyeditabari, Mirsad Hadzikadic, and Wlodek Zadrozny

10:30–11:00 **Morning Coffee Break**

11:00–11:20 A Corpus of Corporate Annual and Social Responsibility Reports: 280 Million Tokens of Balanced Organizational Writing
Sebastian G.M. Händschke, Sven Buechel, Jan Goldenstein, Philipp Poschmann, Tinghui Duan, Peter Walgenbach, and Udo Hahn

11:20–11:40 Word Embeddings-Based Uncertainty Detection in Financial Disclosures
Christoph Kilian Theil, Sanja Stajner, and Heiner Stuckenschmidt

11:40–12:00 A Simple End-to-End Question Answering Model for Product Information
Tuan Lai, Trung Bui, Sheng Li, and Nedim Lipka

12:00–14:00 **Lunch Break**

14:00–14:20 Sentence Classification for Investment Rules Detection
Youness Mansar and Sira Ferradans

14:20–14:40 Leveraging News Sentiment to Improve Microblog Sentiment Classification in the Financial Domain
Tobias Daudert, Paul Buitelaar, and Sapna Negi

14:40–15:00 Implicit and Explicit Aspect Extraction in Financial Microblogs
Thomas Gaillat, Bernardo Stearns, Gopal Sridhar, Ross McDermott, Manel Zarrouk, and Brian Davis

15:00–15:20 Unsupervised Word Influencer Networks from News Streams
Ananth Balashankar, Sunandan Chakraborty, and Lakshminarayanan Subramanian

15:30–16:00 **Afternoon Coffee Break**

16:00–16:30 **Discussion and Wrap-up**

16:30–17:30 **Poster Session - All of the Papers Presented at the Workshop**

RepL4NLP: The Third Workshop on Representation Learning for NLP

Organizers: *Isabelle Augenstein, Kris Cao, He He, Felix Hill, Spandana Gella, Jamie Kiros, Hongyuan Mei, and Dipendra Misra*

Venue: 210

Friday, July 20, 2018

09:30–09:45 **Welcome and Opening Remarks**

09:45–14:45 **Keynote Session**

09:45–10:30 **Invited Talk 1 (Yejin Choi)**

10:30–11:00 **Coffee Break**

11:00–11:45 **Invited Talk 2 (Trevor Cohn)**

11:45–12:30 **Invited Talk 3 (Margaret Mitchell)**

12:30–14:00 **Lunch**

14:00–14:45 **Invited Talk 4 (Yoav Goldberg)**

14:45–15:00 **Outstanding Papers Spotlight Presentations**

15:00–16:30 **Poster Session (including Coffee Break from 15:30-16:00) + Drinks Reception**

- Corpus Specificity in LSA and Word2vec: The Role of Out-of-Domain Documents
  Edgar Altszyler, Mariano Sigman, and Diego Fernandez Slezak
- Hierarchical Convolutional Attention Networks for Text Classification
  Shang Gao, Arvind Ramanathan, and Georgia Tourassi
- Extrofitting: Enriching Word Representation and its Vector Space with Semantic Lexicons
  Hwiyeol Jo and Stanley Jungkyu Choi
- Chat Discrimination for Intelligent Conversational Agents with a Hybrid CNN-LMTGRU Network
  Dennis Singh Moirangthem and Minho Lee
- Text Completion using Context-Integrated Dependency Parsing
  Amr Rekaby Salama, Özge Alacam, and Wolfgang Menzel
- Quantum-Inspired Complex Word Embedding
  Qiuchi Li, Sagar Uprety, Benyou Wang, and Dawei Song
- Natural Language Inference with Definition Embedding Considering Context On the Fly
  Kosuke Nishida, Kyosuke Nishida, Hisako Asano, and Junji Tomita
- Comparison of Representations of Named Entities for Document Classification
  Lidia Pivovarova and Roman Yangarber
- Speeding up Context-based Sentence Representation Learning with Non-autoregressive Convolutional Decoding
  Shuai Tang, Hailin Jin, Chen Fang, Zhaowen Wang, and Virginia de Sa
- Connecting Supervised and Unsupervised Sentence Embeddings
  Gil Levi
- A Hybrid Learning Scheme for Chinese Word Embedding
  Weiguo Sheng and Weiguo Sheng

- Unsupervised Random Walk Sentence Embeddings: A Strong but Simple Baseline
  Kawin Ethayarajh and Graeme Hirst
- Evaluating Word Embeddings in Multi-label Classification Using Fine-Grained Name Typing
  Yadollah Yaghoobzadeh, Katharina Kann, and Hinrich Schütze
- A Dense Vector Representation for Open-Domain Relation Tuples
  Ade Romadhony, Alfan Farizki Wicaksono, Ayu Purwarianti, and Dwi Hendratmo Widyantoro
- Exploiting Common Characters in Chinese and Japanese to Learn Cross-Lingual Word Embeddings via Matrix Factorization
  Jilei Wang, Shiying Luo, Weiyan Shi, Tao Dai, and Shu-Tao Xia
- WordNet Embeddings
  Chakaveh Saedi, António Branco, João António Rodrigues, and João Silva
- Knowledge Graph Embedding with Numeric Attributes of Entities
  Yanrong Wu and Zhichun Wang
- Injecting Lexical Contrast into Word Vectors by Guiding Vector Space Specialisation
  Ivan Vulić
- Characters or Morphemes: How to Represent Words?
  Ahmet Üstün, Murathan Kurfalı, and Burcu Can
- Learning Hierarchical Structures On-The-Fly with a Recurrent-Recursive Model for Sequences
  Athul Paul Jacob, Zhouhan Lin, Alessandro Sordoni, and Yoshua Bengio
- Limitations of Cross-Lingual Learning from Image Search
  Mareike Hartmann and Anders Søgaard
- Learning Semantic Textual Similarity from Conversations
  Yinfei Yang, Steve Yuan, Daniel Cer, Sheng-Yi Kong, Noah Constant, Petr Pilar, Heming Ge, Yun-hsuan Sung, Brian Strope, and Ray Kurzweil
- Multilingual Seq2seq Training with Similarity Loss for Cross-Lingual Document Classification
  Katherine Yu, Haoran Li, and Barlas Oguz
- LSTMs Exploit Linguistic Attributes of Data
  Nelson F. Liu, Omer Levy, Roy Schwartz, Chenhao Tan, and Noah A. Smith
- Learning Distributional Token Representations from Visual Features
  Samuel Broscheit
- Jointly Embedding Entities and Text with Distant Supervision
  Denis Newman-Griffis, Albert M. Lai, and Eric Fosler-Lussier
- A Sequence-to-Sequence Model for Semantic Role Labeling
  Angel Daza and Anette Frank
- Predicting Concreteness and Imageability of Words Within and Across Languages via Word Embeddings
  Nikola Ljubešić, Darja Fišer, and Anita Peti-Stantić

16:30–17:30 **Panel Discussion**

17:30–17:40 **Closing Remarks + Best Paper Awards Announcement**

Challenge-MML: The First Grand Challenge and Workshop on Human Multimodal Language

Organizers: *Amir Zadeh, Paul Pu Liang, Louis-Philippe Morency, Soujanya Poria, Erik Cambria, and Stefan Scherer*

Venue: 217

Friday, July 20, 2018

09:00–10:30 **Session 1**

09:00–09:10 **Opening Remarks**

09:10–10:00 **Keynote (Bing Liu)**

10:00–10:10 Getting the subtext without the text: Scalable multimodal sentiment classification from visual and acoustic modalities
Nathaniel Blanchard, Daniel Moreira, Aparna Bharati, and Walter Scheirer

10:10–10:20 Recognizing Emotions in Video Using Multimodal DNN Feature Fusion
Jennifer Williams, Steven Kleinegesse, Ramona Comanescu, and Oana Radu

10:20–10:30 Multimodal Relational Tensor Network for Sentiment and Emotion Classification
Saurav Sahay, Shachi H Kumar, Rui Xia, Jonathan Huang, and Lama Nachman

10:30–11:00 **Coffee Break**

11:00–12:30 **Session 2**

11:00–11:50 **Keynote (Sharon Oviatt)**

11:50–12:00 **Advances in Multimodal Datasets (Paul Pu Liang)**

12:00–12:10 Convolutional Attention Networks for Multimodal Emotion Recognition from Speech and Text Data
Woo Yong Choi, Kyu Ye Song, and Chan Woo Lee

12:10–12:20 Sentiment Analysis using Imperfect Views from Spoken Language and Acoustic Modalities
Imran Sheikh, Sri Harsha Dumpala, Rupayan Chakraborty, and Sunil Kumar Kopparapu

12:20–12:30 Polarity and Intensity: the Two Aspects of Sentiment Analysis
Leimin Tian, Catherine Lai, and Johanna Moore

12:30–13:30 **Lunch Break**

13:30–15:00 **Session 3**

13:30–14:20 **Keynote (Roland Goecke)**

14:20–14:30 ASR-based Features for Emotion Recognition: A Transfer Learning Approach
Noé Tits, Kevin El Haddad, and Thierry Dutoit

14:30–14:40 Seq2Seq2Sentiment: Multimodal Sequence to Sequence Models for Sentiment Analysis
Hai Pham, Thomas Manzini, Paul Pu Liang, and Barnabas Poczos

14:40–14:50 DNN Multimodal Fusion Techniques for Predicting Video Sentiment
Jennifer Williams, Ramona Comanescu, Oana Radu, and Leimin Tian

14:50–15:00 **Grand Challenge Results**

15:00–15:05 **Workshop End**

SocialNLP: The Sixth Workshop on Natural Language Processing for Social Media

Organizers: *Lun-Wei Ku and Cheng-Te Li*
Venue: 216

Friday, July 20, 2018

09:20–10:30 **Keynote Speech (I):The Search for Emotions, Creativity, and Fairness in Language (Dr. Saif Mohammad (NSF))**

10:30–11:00 **Coffee Break**

11:00–12:20 **Technical Session 1**

- Sociolinguistic Corpus of WhatsApp Chats in Spanish among College Students
  Alejandro Dorantes, Gerardo Sierra, Tlauhlia Yamín Donohue Pérez, Gemma Bel-Enguix, and Mónica Jasso Rosales
- A Crowd-Annotated Spanish Corpus for Humor Analysis
  Santiago Castro, Luis Chiruzzo, Aiala Rosá, Diego Garat, and Guillermo Moncecchi
- A Twitter Corpus for Hindi-English Code Mixed POS Tagging
  Kushagra Singh, Indira Sen, and Ponnurangam Kumaraguru
- Detecting Offensive Tweets in Hindi-English Code-Switched Language
  Puneet Mathur, Rajiv Shah, Ramit Sawhney, and Debanjan Mahata

12:20–13:20 **Lunch**

13:20–14:30 **Keynote Speech (II): Understanding Online Social Behaviors through Automatic Language Analysis (Dr. Yi-Chia Wang (Uber))**

14:30–15:30 **EmotionX Challenge Session**

- SocialNLP 2018 EmotionX Challenge Overview: Recognizing Emotions in Dialogues
  Chao-Chun Hsu and Lun-Wei Ku
- EmotionX-DLC: Self-Attentive BiLSTM for Detecting Sequential Emotions in Dialogues
  Linkai Luo, Haiqin Yang, and Francis Y. L. Chin
- EmotionX-AR: CNN-DCNN autoencoder based Emotion Classifier
  Sopan Khosla
- EmotionX-SmartDubai_NLP: Detecting User Emotions In Social Media Text
  Hessa AlBalooshi, Shahram Rahmanian, and Rahul Venkatesh Kumar
- EmotionX-Area66: Predicting Emotions in Dialogues using Hierarchical Attention Network with Sequence Labeling
  Rohit Saxena, Savita Bhat, and Niranjan Pedanekar
- EmotionX-JTML: Detecting emotions with Attention
  Johnny Torres

15:30–16:00 **Coffee Break**

16:00–17:00 **Technical Session 2**

- Towards Automation of Sense-type Identification of Verbs in OntoSenseNet
  Sreekavitha Parupalli, Vijjini Anvesh Rao, and Radhika Mamidi
- Improving Classification of Twitter Behavior During Hurricane Events
  Kevin Stowe, Jennings Anderson, Martha Palmer, Leysia Palen, and Ken Anderson

- Political discourse classification in social networks using context sensitive convolutional neural networks
  Aritz Bilbao-Jayo and Aitor Almeida

17:00–17:10 **Closing**

NMT: The 2nd Workshop on Neural Machine Translation and Generation

Organizers: *Alexandra Birch, Andrew Finch, Thang Luong, Graham Neubig, and Yusuke Oda*
Venue: 211

Friday, July 20, 2018

09:00–09:10 **Welcome and Opening Remarks**

- Findings of the Second Workshop on Neural Machine Translation and Generation
  Alexandra Birch, Andrew Finch, Minh-Thang Luong, Graham Neubig, and Yusuke Oda

09:10–10:00 **Keynote 1 (Jacob Devlin)**

10:00–10:30 **Shared Task Overview**

10:30–11:00 **Coffee Break**

11:00–11:30 **Marian: Fast Neural Machine Translation in C++**

11:30–12:20 **Keynote 2 (Rico Sennrich)**

12:20–13:20 **Lunch Break**

13:20–13:50 **Best Paper Session**

13:50–14:40 **Keynote 3 (Jason Weston)**

14:40–15:30 **Keynote 4 (Yulia Tsvetkov)**

15:30–16:00 **Coffee Break**

16:00–17:30 **Poster Session**

- A Shared Attention Mechanism for Interpretation of Neural Automatic Post-Editing Systems
  Inigo Jauregi Unanue, Ehsan Zare Borzeshi, and Massimo Piccardi
- Iterative Back-Translation for Neural Machine Translation
  Vu Cong Duy Hoang, Philipp Koehn, Gholamreza Haffari, and Trevor Cohn
- Inducing Grammars with and for Neural Machine Translation
  Yonatan Bisk and Ke Tran
- Regularized Training Objective for Continued Training for Domain Adaptation in Neural Machine Translation
  Huda Khayrallah, Brian Thompson, Kevin Duh, and Philipp Koehn
- Controllable Abstractive Summarization
  Angela Fan, David Grangier, and Michael Auli
- Enhancement of Encoder and Attention Using Target Monolingual Corpora in Neural Machine Translation
  Kenji Imamura, Atsushi Fujita, and Eiichiro Sumita
- Document-Level Adaptation for Neural Machine Translation
  Sachith Sri Ram Kothur, Rebecca Knowles, and Philipp Koehn
- On the Impact of Various Types of Noise on Neural Machine Translation
  Huda Khayrallah and Philipp Koehn
- Bi-Directional Neural Machine Translation with Synthetic Parallel Data
  Xing Niu, Michael Denkowski, and Marine Carpuat

- Multi-Source Neural Machine Translation with Missing Data
  Yuta Nishimura, Katsuhito Sudoh, Graham Neubig, and Satoshi Nakamura
- Towards one-shot learning for rare-word translation with external experts
  Ngoc-Quan Pham, Jan Niehues, and Alexander Waibel
- NICT Self-Training Approach to Neural Machine Translation at NMT-2018
  Kenji Imamura and Eiichiro Sumita
- Fast Neural Machine Translation Implementation
  Hieu Hoang, Tomasz Dwojak, Rihards Krislauks, Daniel Torregrosa, and Kenneth Heafield
- OpenNMT System Description for WNMT 2018: 800 words/sec on a single-core CPU
  Jean Senellart, Dakun Zhang, Bo Wang, Guillaume Klein, Jean-Pierre Ramatchandirin, Josep Crego, and Alexander Rush
- Marian: Cost-effective High-Quality Neural Machine Translation in C++
  Marcin Junczys-Dowmunt, Kenneth Heafield, Hieu Hoang, Roman Grundkiewicz, and Anthony Aue
- On Individual Neurons in Neural Machine Translation
  D. Anthony Bau, Yonatan Belinkov, Hassan Sajjad, Nadir Durrani, Fahim Dalvi, and James Glass
- Parameter Sharing Strategies in Neural Machine Translation
  Sébastien Jean, Stanislas Lauly, and Kyunghyun Cho
- Modeling Latent Sentence Structure in Neural Machine Translation
  Joost Bastings, Wilker Aziz, Ivan Titov, and Khalil Simaan
- Extreme Adaptation for Personalized Neural Machine Translation
  Paul Michel and Graham Neubig
- Exploiting Semantics in Neural Machine Translation with Graph Convolutional Networks
  Diego Marcheggiani, Joost Bastings, and Ivan Titov

17:30–17:40 **Closing Remarks**

NEWS: The Seventh Named Entities Workshop

Organizers: *Nancy Chen, Rafael E. Banchs, Xiangyu Duan, Min Zhang, and Haizhou Li*

Venue: 208

Friday, July 20, 2018

08:30–08:40 **Opening Remarks**

08:40–09:00 Automatic Extraction of Entities and Relation from Legal Documents
Judith Jeyafreeda Andrew

09:00–09:20 Connecting Distant Entities with Induction through Conditional Random Fields for Named Entity Recognition: Precursor-Induced CRF
Wangjin Lee and Jinwook Choi

09:20–09:40 A Sequence Learning Method for Domain-Specific Entity Linking
Emrah Inan and Oguz Dikenelli

09:40–10:00 Attention-based Semantic Priming for Slot-filling
Jiewen Wu, Rafael E. Banchs, Luis Fernando D'Haro, Pavitra Krishnaswamy, and Nancy Chen

10:00–10:20 Named Entity Recognition for Hindi-English Code-Mixed Social Media Text
Vinay Singh, Deepanshu Vijay, Syed Sarfaraz Akhtar, and Manish Shrivastava

10:30–11:00 **Coffee Break**

11:00–11:20 Forms of Anaphoric Reference to Organisational Named Entities: Hoping to widen appeal, they diversified
Christian Hardmeier, Luca Bevacqua, Sharid Loáiciga, and Hannah Rohde

11:20–11:40 Named-Entity Tagging and Domain adaptation for Better Customized Translation
Zhongwei Li, Xuancong Wang, AiTi Aw, Eng Siong Chng, and Haizhou Li

12:00–14:00 **Lunch**

14:10–14:20 NEWS 2018 Whitepaper
Nancy Chen, Xiangyu Duan, Min Zhang, Rafael E. Banchs, and Haizhou Li

14:20–14:40 Report of NEWS 2018 Named Entity Transliteration Shared Task
Nancy Chen, Rafael E. Banchs, Min Zhang, Xiangyu Duan, and Haizhou Li

14:40–15:00 Statistical Machine Transliteration Baselines for NEWS 2018
Snigdha Singhania, Minh Nguyen, Gia H Ngo, and Nancy Chen

15:00–15:20 A Deep Learning Based Approach to Transliteration
Soumyadeep Kundu, Sayantan Paul, and Santanu Pal

15:30–16:00 **Coffee Break**

16:00–16:20 Comparison of Assorted Models for Transliteration
Saeed Najafi, Bradley Hauer, Rashed Rubby Riyadh, Leyuan Yu, and Grzegorz Kondrak

16:20–16:40 Neural Machine Translation Techniques for Named Entity Transliteration
Roman Grundkiewicz and Kenneth Heafield

16:40–17:00 Low-Resource Machine Transliteration Using Recurrent Neural Networks of Asian Languages
Ngoc Tan Le and Fatiha Sadat

NLPOSS: Workshop for NLP Open Source Software

Organizers: *Eunjeong L. Park, Masato Hagiwara, Dmitrijs Milajevs, and Liling Tan*
Venue: 209

Friday, July 20, 2018

08:45–09:00 **Loading Presentations to Computer**

09:00–09:05 **Opening Remarks**

09:05–09:50 **Invited Talk 1 (Joel Nothman)**

09:50–10:30 **Lightning Presentation for Posters Session 1**

10:30–11:00 **Coffee Break**

11:00–11:45 **Poster Session 1**

- AllenNLP: A Deep Semantic Natural Language Processing Platform
  Matt Gardner, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson F. Liu, Matthew Peters, Michael Schmitz, and Luke Zettlemoyer
- Stop Word Lists in Free Open-source Software Packages
  Joel Nothman, Hanmin Qin, and Roman Yurchak
- Texar: A Modularized, Versatile, and Extensible Toolbox for Text Generation
  Zhiting Hu, Zichao Yang, Tiancheng Zhao, Haoran Shi, Junxian He, Di Wang, Xuezhe Ma, Zhengzhong Liu, Xiaodan Liang, Lianhui Qin, Devendra Singh Chaplot, Bowen Tan, Xingjiang Yu, and Eric Xing
- The ACL Anthology: Current State and Future Directions
  Daniel Gildea, Min-Yen Kan, Nitin Madnani, Christoph Teichmann, and Martin Villalba
- The risk of sub-optimal use of Open Source NLP Software: UKB is inadvertently state-of-the-art in knowledge-based WSD
  Eneko Agirre, Oier Lopez de Lacalle, and Aitor Soroa

12:00–14:00 **Lunch**

14:00–14:45 **Invited Talk 2 (Christopher Manning)**

14:45–15:30 **Lightning Presentation for Posters 2**

15:30–16:00 **Break**

16:00–16:45 **Poster Session 2**

- Baseline: A Library for Rapid Modeling, Experimentation and Development of Deep Learning Algorithms targeting NLP
  Daniel Pressel, Sagnik Ray Choudhury, Brian Lester, Yanjie Zhao, and Matt Barta
- OpenSeq2Seq: Extensible Toolkit for Distributed and Mixed Precision Training of Sequence-to-Sequence Models
  Oleksii Kuchaiev, Boris Ginsburg, Igor Gitman, Vitaly Lavrukhin, Carl Case, and Paulius Micikevicius
- Integrating Multiple NLP Technologies into an Open-source Platform for Multilingual Media Monitoring
  Ulrich Germann, Renars Liepins, Didzis Gosko, and Guntis Barzdins
- The Annotated Transformer
  Alexander Rush

16:45–17:30 **Invited Talk 3 (Matthew Honnibal and Ines Montani)**

17:30–17:35 **Closing Remarks**

7

Anti-harassment Policy

The open exchange of ideas, the freedom of thought and expression, and respectful scientific debate are central to the aims and goals of a ACL conference. These require a community and an environment that recognizes the inherent worth of every person and group, that fosters dignity, understanding, and mutual respect, and that embraces diversity. For these reasons, ACL is dedicated to providing a harassment-free experience for participants at our events and in our programs.

Harassment and hostile behavior are unwelcome at any ACL conference. This includes: speech or behavior (including in public presentations and on-line discourse) that intimidates, creates discomfort, or interferes with a person's participation or opportunity for participation in the conference. We aim for ACL conferences to be an environment where harassment in any form does not happen, including but not limited to: harassment based on race, gender, religion, age, color, national origin, ancestry, disability, sexual orientation, or gender identity. Harassment includes degrading verbal comments, deliberate intimidation, stalking, harassing photography or recording, inappropriate physical contact, and unwelcome sexual attention.

It is the responsibility of the community as a whole to promote an inclusive and positive environment for our scholarly activities. In addition, any participant who experiences harassment or hostile behavior may contact any current member of the ACL Board or contact Priscilla Rasmussen, who is usually available at the registration desk of the conference. Please be assured that if you approach us, your concerns will be kept in strict confidence, and we will consult with you on any actions taken.

The ACL board members are listed at:

`https://www.aclweb.org/portal/about`

The full policy and its implementation is defined at:

`https://www.aclweb.org/adminwiki/index.php?title=Anti-Harassment_Policy`

8

ACL Author Guidelines

Preserving Double Blind Review

The following rules and guidelines are meant to protect the integrity of double-blind review and ensure that submissions are reviewed fairly. The rules make reference to the anonymity period, which runs from 1 month before the submission deadline up to the date when your paper is either accepted, rejected, or withdrawn.

- You may not make a non-anonymized version of your paper available online to the general community (for example, via a preprint server) during the anonymity period. By a version of a paper we understand another paper having essentially the same scientific content but possibly differing in minor details (including title and structure) and/or in length (e.g., an abstract is a version of the paper that it summarizes).

- If you have posted a non-anonymized version of your paper online before the start of the anonymity period, you may submit an anonymized version to the conference. The submitted version must not refer to the non-anonymized version, and you must inform the program chair(s) that a non-anonymized version exists. You may not update the non-anonymized version during the anonymity period, and we ask you not to advertise it on social media or take other actions that would further compromise double-blind reviewing during the anonymity period.

- Note that, while you are not prohibited from making a non-anonymous version available online before the start of the anonymity period, this does make double-blind reviewing more difficult to maintain, and we therefore encourage you to wait until the end of the anonymity period if possible. Alternatively, you may consider submitting your work to the Computational Linguistics journal, which does not require anonymization and has a track for "short" (i.e., conference-length) papers.

Citation and Comparison

If you are aware of previous research that appears sound and is relevant to your work, you should cite it even if it has not been peer-reviewed, and certainly if it influenced your own work. However, refereed publications take priority over unpublished work reported in preprints. Specifically:

- You are expected to cite all refereed publications relevant to your submission, but you may be excused for not knowing about all unpublished work (especially work that has been recently posted and/or is not widely cited).
- In cases where a preprint has been superseded by a refereed publication, the refereed publication should be cited in addition to or instead of the preprint version.

Papers (whether refereed or not) appearing less than 3 months before the submission deadline are considered contemporaneous to your submission, and you are therefore not obliged to make detailed comparisons that require additional experimentation and/or in-depth analysis.

9 Local Guide

This guide was originally written by Tony Wirth, with additions by Jey Han Lau, Trevor Cohn and Timothy Baldwin.

Introduction

Melburnians are only too happy to tell you that their city is the cultural capital of Australia, or the sporting capital of Australia. Though true, Melbourne's real international claim is that its food is world class and its coffee is even better, with outstanding cafés and coffee roasters scattered across the city. A combination of strong competition and a multicultural population that frequently dines out has kindled a vibrant café and restaurant scene.

The CBD, as we like to call it, aka "the city" or Melbourne 3000 – or Downtown if you want to sound like a tourist – has the highest density of restaurants in greater Melbourne. In general, the streets to the east of Elizabeth Street, especially, east of Russell Street, are best for evenings. For an outstanding gelato experience, with experimental flavours, drop into Gelateria Primavera. Chinatown has its spine along Little Bourke Street, including the long-standing Golden Orchids.

Closest to the conference venue, the most vibrant area is near the corner of Katherine Place and Flinders Lane. Further east along Flinders Lane, you can find Dukes coffee roasters. Degraves Street, either side of Flinders Lane, is packed with cafés, and is very popular with visitors and locals alike.

There are several quality *rooftop bars* in the CBD, including Bomba, Siglo, Campari House, Rooftop Bar (Curtin House), Madame Brussels, and Red Hummingbird. For a view of the city, and of local commuters rushing for their trains, head to the delightful Arbory. Either side, along the Yarra River, Ponyfish Island and Riverland are gems.

Explore Melbourne's *laneways*: some of them have amazing places to eat (Hardware Lane, Centre Place, Degraves Street), some have fantastic artwork (Hosier Lane), some are just a little scary looking.

To browse a list of curated cafés, restaurants and other notable places interactively on your smartphone, please check out the following Google Maps Places: `https://goo.gl/maps/6Q96QDCHbJD2`.

Useful Links

Food https://www.broadsheet.com.au/melbourne

Cafés https://www.beanhunter.com/melbourne
https://www.broadsheet.com.au/melbourne/guides/best-coffee

Things To Do http://www.visitvictoria.com/Regions/Melbourne

Getting Around

Melbourne is famous for its trams, and has one of the most extensive tram networks in the world. The CBD is designated a Free Tram Zone, meaning that you can ride any tram for free, although beware, as the Free Tram Zone finishes one stop short of the conference venue and ticket inspectors frequent the fringes of the zone trying to catch out those who ride without a ticket, and are infamous for their intolerance (including tourists). If you wish to catch a tram beyond the Free Tram Zone, you will need to purchase a Myki ticket from a newsagent or one of the many vending machines at tram stops, and "touch on" each time you get on a tram. Note that if you touch on within the Free Tram Zone, you will be charged. No, not the most user-friendly system in the world, but there are clear announcements in the trams of whether you are in the Free Tram Zone or not. The same Myki ticket will work on local trains and busses, should you need to travel further afield. Google maps gives a good overview of public transport options, as does https://www.ptv.vic.gov.au/ and the PTV phone app.

Uber operates in Melbourne, and tends to be cheaper than taxis for short trips.

Surrounding Suburbs

South Melbourne Some of the great cafés of Melbourne, St Ali and Chez Dré, are just a few blocks south of the conference venue.

Fitzroy Taking the Number 96 tram north for about 25 minutes, you'll arrive in Fitzroy. Apart from the CBD, this inner suburb has the highest concentration of restaurants. There are classic atmospheric pubs, such as The Napier, Labour in Vain, and The Standard, as well as the newer Naked for Satan with its rooftop Naked in the Sky. The cafés are top notch, with Industry Beans featuring in an article by *The Huffington Post* on the most hipster neighborhoods in the world. In general, Brunswick Street and Smith Street reward aimless wandering, and can be reached on the 59 tram.

Bayside Port Melbourne is a short ride down the 109 tram. Bay Street has several good eateries, including bakery Noisette. St Kilda is the southern terminus of the Number 96 tram. The Sunday esplanade market is a Melbourne classic, the scenic railway at Luna Park is the oldest continuously running roller coaster in the world, and the continental cakes along Acland Street have kept the area buzzing for decades.

Guide to be a Tourist in Melbourne

- Buying coffee at Starbucks (Melburnians do pride themselves on their local café culture!)
- Meaning anything other than "Aussie rules football" when referring to "football" (Melburnians are, in large part, famously one-eyed when it comes to football codes)... in fact using the term "football" at all, as Australians love to abbreviate everything, including "footie"

- Not having an immediate response and breaking into impassioned dialogue/song when asked "who do you barrack for" (referring, of course, to the footie team you support)

- Tipping — tipping culture is very limited in Melbourne, and it is only at high-end restaurants where there is really any expectation of a tip, and even here it is optional. When eating out in large groups, high-end restaurants will sometimes charge a group surcharge, meaning even less reason to tip. Certainly there is no need to tip in taxis or at cafés (other than in the form of loose change in the tipping jar).

- Not having a humorous/whimsical come-back at the ready at all times — Melburnians are generally a very friendly, laid-back bunch who try not to take themselves too seriously (except when it comes to footie, of course), and like to light-heartedly "take the piss" when the opportunity arises

- Comparing Sydney with Melbourne favourably in any way — Melburnians are very proud of their city, and fiercely territorial when it comes to comparisons with Sydney

- Asking about "Australian" eating options — Melbourne is proudly multicultural and very proud of its "foodie" culture, and "Australian" cuisine is representative of that: a melting pot of the myriad of different cuisines of the many migrant groups who make up its population; if there is an "Australian" cuisine, it is in the blending/fusion of different cuisines. Particular cuisines where Melbourne excels include Chinese, Japanese, Korean, Vietnamese, Malaysian, Indonesian, Italian, and Greek, with many fantastic options within easy access of the conference venue.

- Being suspicious of off-laneway, hidden-by-dumpsters cafés — in general, the more obscure the location, the more hipster the café and more authentically local the clientele.

- Walking/riding/driving on the right — yes, we get that much of the rest of the world drives on the left, but we don't.

- Not holding the door for the person behind you — not particular to Melbourne necessarily, and people aren't being chauvinistic or patronising in any way, it's just a politeness thing here.

- Getting annoyed when people approach you to ask if you need directions when you have a map out or are otherwise looking lost — don't feel patronised or defensive, it's just the locals being friendly.

- Asking where the kangaroos or koalas are in central Melbourne — kangaroos (or "roos" as the locals call them) outnumber people in Australia (koalas are less common), yes, but you won't find them hopping down Swanston St. Where you will spot them is in places where there is open farmland, where they thrive.

- Being offended at the standards of "colourful" language in Australia — it's not uncommon for Australians to pepper their speech with swear words, as a way of colouring their speech, with no intention of offending anyone's sensibilities.

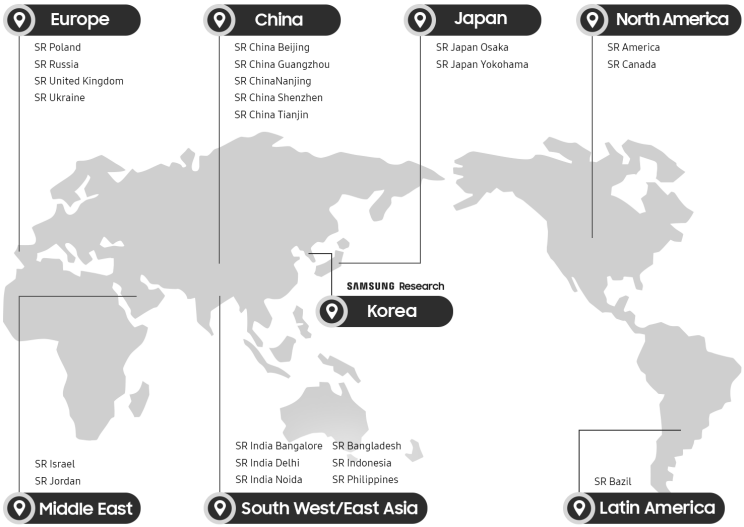
Index

SAMSUNG Research

Shape the Future with Innovation and Intelligence

To create a valuable future through innovationn and intelligence

Global R&D Network



RESEARCH AREAS

Artificial Intelligence | Data Intelligence | IoT | Tizen | Smart Machine | Security | Next Generation Media

HOW TO APPLY

- career.sr@samsung.com
- http://research.samsung.com

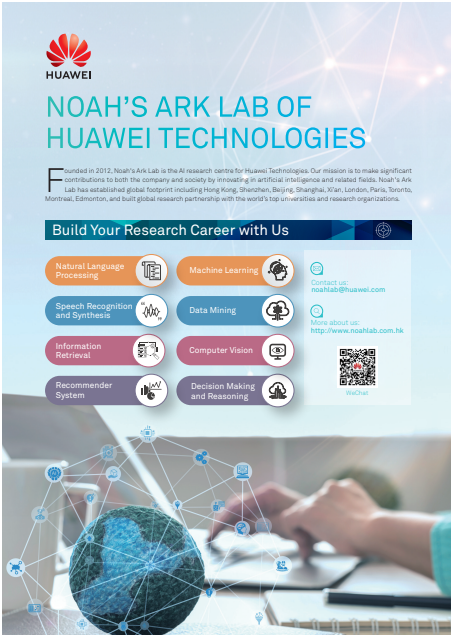
Megagon Labs

Megagon Labs (formerly Recruit Institute of Technology) is the technology hub and research lab for Recruit Holdings, a company that provides over 200 online services in the areas of human resources, travel, housing, education, restaurants and many other areas in which people make daily lifestyle decisions.

We conduct world-class research in the several areas, including natural language processing, machine learning, data management, data integration, and artificial intelligence. We collaborate with research in universities and publish in top-notch conferences.

Megagon Labs is hiring Research Scientists and Software Engineers. Visit our website at www.megagon.ai for more information.


Tencent
AI
Tencent is applying its technology leadership, deep understanding of connection and user behavior and open platform/ecosystem to create AI breakthroughs that can solve the world's most pressing challenges. We are not innovating for its own sake or simply seeking to win AI contents. We are exploring how AI can enhance the quality of human life.
Tencent's mission is to enhance the quality of human life through Internet services and AI enables us to do this in powerful new ways. We are focused on solving key problems for our users, customers and society. Through fundamental and applied research, we seek to create products and services that make the benefits of AI accessible to everyone, everywhere.
Tencent AI Job List
Principle/Senior Researcher-Natural Language & Speech Processing, Tencent AI Lab
Location: Seattle Bellevue, United States
Research Scientist-Natural Language Processing, Tencent Lingo Lab
Location: Beijing / Shenzhen / Chengdu, China
Senior Researcher-Natural Language Processing , Conversational Bots, Tencent Cloud
Location: Shenzhen, China
(Senior) Researcher-Natural Language Processing, Wechat
Location: Beijing/Guangzhou/Shenzhen, China
Senior Researcher-Natural Language & Speech Processing, Tencent Games
Location: Shenzhen, China/Los Angeles, United States
Senior Researcher-Natural Language Processing, Financial Data Application Department
Location: Shenzhen, China
Connect with us
LinkedIn:@tencent
WeChat:@QQjobs
Tencent


HUAWEI
NOAH'S ARK LAB OF HUAWEI TECHNOLOGIES
Founded in 2012, Noah's Ark Lab is the AI research centre for Huawei Technologies. Our mission is to make significant contributions to both the company and society by innovating in artificial intelligence and related fields. Noah's Ark Lab has established global footprint including Hong Kong, Shenzhen, Beijing, Shanghai, Xi'an, London, Paris, Toronto, Montreal, Edmonton, and built global research partnership with the world's top universities and research organizations.
Build Your Research Career with Us
Natural Language Processing
Machine Learning
Speech Recognition and Synthesis
Data Mining
Information Retrieval
Computer Vision
Recommender System
Decision Making and Reasoning
Contact us:
noahlab@huawei.com
More about us:
http://www.noahlab.com.hk
WeChat

Getting from Conference to Student Recruitment and Social Event

Precinct Map

Convention Centre Ground Floor

Convention Centre Level Two

ACL gratefully acknowledges the following sponsors for their support!

Diamond

Google

facebook

ByteDance 字节跳动

SAMSUNG

Platinum

amazon

Megagon Labs

Baidu 百度

京东

Tencent 腾讯

Gold

Microsoft

IBM Research AI

CVTE 视源股份

NAVER | LINE

digital health crc

Silver

NUANCE

UBER

HUAWEI

duolingo

ELSEVIER

Bronze

SEP California 2013 ISI NLP PROTECT OUR COAST & OCEAN

Australian Government
Department of Defence
Science and Technology

Student Volunteer Sponsors

DATA 61

CSIRO

Supporter

MELBOURNE Victoria Australia

ROAM

SAP

MELBOURNE Convention Bureau

SINTELIX