CONFERENCE PROCEEDINGS

SEPTEMBER 17-21

EMNLP 2015

CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING

LISBON

contact@emnlp2015.org

WWW.EMNLP2015.ORG

PLATINUM SPONSORS

Bloomberg Baidu 百度

GOLD SPONSORS

Google facebook LinkedIn

SILVER SPONSORS

RECRUIT Recruit Technologies Co., Ltd. NUANCE amazon AI @ ISI VoiceBox TECHNOLOGIES DigitalGenius Artificial Intelligence for Brands

EMNLP 2015 gratefully acknowledges the following sponsors for their support:

Platinum

Gold

Silver

Bronze

Best Student Paper

Supporters

Institutional partners

Official Airline

Order print-on-demand copies from:

Curran Associates
57 Morehouse Lane
Red Hook, New York 12571 USA
Tel: +1-845-758-0400
Fax: +1-845-758-2633
curran@proceedings.com

©2015 The Association for Computational Linguistics
ISBN: 978-1-941643-32-7

Table of Contents

Preface by the General Chair

August 25, 2015

Welcome to the 2015 Conference on Empirical Methods in Natural Language Processing. EMNLP is annually organized by SIGDAT, the Association for Computational Linguistics' special interest group on linguistic data and corpus-based approaches to NLP. This year the conference will be held on September 17–21 in the enchanting city of Lisbon, Portugal.[1]

EMNLP has continued to increase in prominence as one of the most important conferences in Natural Language Processing (NLP). This year the conference has experienced an unprecedented boost in submitted papers. I believe that this reflects both the growth of the NLP field and also the health and strength of the conference itself, with a history of many years of solid work. With this level of interest at submission time, we are also expecting a record attendance. The conference will span a five-day period this year, and it requires a growing organization structure.

Some of the features introduced in EMNLP 2014 will continue this year (e.g., tutorials, new chairs, posters as parallel sessions, flat rates and flexibility for tutorials and workshops, etc.). We also introduce some innovations, like a revised selection process for which talks are presented as talks versus posters.

This year I had the privilege of coordinating the conference from my General Chair position. This has been a very instructive and enriching exercise which showed me the conference as a whole, from many different angles. Prefaces in the proceedings invariably praise the team of organizers. This one will not be an exception. Organizing a large conference as EMNLP requires excellent people working as a team in multiple interrelated tasks. I have been lucky to work with an outstanding team of people, from whom I learnt a lot. These aren't empty words. I would like to thank each and every chair for the hard work that made the conference a reality.

The Program Chairs, Jian Su and Chris Callison-Burch, did an excellent job at putting together a very interesting program with over 300 papers. They had to deal with a very large number of submissions, which exceeded even our most optimistic expectations. As a consequence, they were forced to be creative and to find solutions on the fly to adapt to the situation. They recruited the largest ever program committee and successfully managed a huge reviewing and decision making process under a very tight schedule. A real gift for the general chair. They complemented the program with very interesting keynote speakers, Yoshua Bengio and Justin Grimmer who will present exciting research topics for our community.

The EMNLP 2015 main conference is accompanied by 7 workshops and 8 tutorials during the first two days. The Workshops Chairs, Zornitsa Kozareva and Jörg Tiedemann, and the Tutorials Chairs, Maggie Li and Khalil Sima'an, conducted the selection processes in a joint effort with the other ACL conferences in 2015 (NAACL and ACL-IJCNLP). This has been the standard procedure from last years. It has the advantage of starting early, avoiding duplicated reviewing and allowing a more balanced selection among conferences. EMNLP attracted a varied and interesting set of workshops and tutorials, which gives more value to the conference.

Daniele Pighin and Yuval Marton were responsible for the always difficult and sometimes thankless task of putting together the conference publications. This is a very complex effort which involves coordination with almost everyone in the team under the pressure of hard publication deadlines. Yuval is serving in this position for a second year. Staggered two year terms for publication chairs is a new addition for

[1] Conference website: `http://www.emnlp2015.org`

EMNLP starting this year, and we hope that it will be a permanent feature. In the first year, publication chairs will learn and do the bulk of proceedings compilation. During the second year their role will be more advisory, instructing and helping the first-year chair. This procedure will help the transmission of the necessary know-how from year to year. Thanks to Yuval and Daniele for accepting the challenge and making it work wonderfully. Finally, this is the second year that EMNLP uses a mobile app for the conference program (Conference4me). The publication chairs also coordinated the integration of the app with SoftConf, which is now smoother and more seamless.

The local organization team was led by André Martins and João Graça. They did an amazing job, working hard and with all the complexities and subtleties of local arrangements. One of the keys for the success was the creation of a large team of local organizers with clearly defined roles and responsibilities. They appointed very committed people: Isabel Trancoso (Local Publicity Chair), Fernando Batista (Handbook Chair), Bruno Martins (Website and App Chair), Luísa Coheur (Student Volunteer Coordinator), and Helena Moniz (Local Arrangements Chair). Thanks to all. I am especially pleased about the new website, which was revamped and looks more professional everyday. This is certainly a good investment for the future.

A large conference as EMNLP needs to focus on dissemination activities too. Barbara Plank acted as the international Publicity Chair. She did a fantastic job and coordinated very well with the local publicity and the website chairs. The calls for papers, calls for participation, and main news of the conference were timely distributed through ACL, the usual distribution lists, and also through the conference website and two Facebook and Twitter accounts. The EMNLP2015 Twitter account garnered more followers than in previous years.

I am really grateful to SIGDAT. Its secretary, Noah Smith, acted as the liasion between SIGDAT and the conference organizers. He was always available and ready to help with very good advice. SIGDAT also provided the funds for the student scholarship program. These grants help covering traveling expenses to a dozen of students. The committee appointed for collecting the applications and making the decisions was formed by Francisco Guzmán and Lluís Padró, who had to analyze all the information and decide the awardees in only a few days.

Another sign of the health of EMNLP and the field in general is the interest showed by sponsors. Thanks to the work of our sponsorship team, formed by João Graça and Hang Li, in coordination with the ACL International Sponsorship Committee, we got a record number of 13 sponsors for EMNLP 2015 (2 platinum, 3 gold, 6 silver and 2 bronze). In addition to these direct sponsors, we also have several smaller supporters, exhibitors, and institutional partners. We are extremely grateful to all these companies and institutions, which make a better conference possible at a more affordable registration fee.

Additionally, we counted on the invaluable help of Priscilla Rasmussen, supporting the local organization in all fronts with her broad experience. She took care of the registration process too. We also got very good advice, know-how, and helpful software and forms from last year general chair and local organizers, Alessandro Moschitti and Kareem Darwish. Thank you.

Finally, I would like to thank the authors of submitted and accepted papers, and all the attendees to the conference, who will be the main actors from September 17 to September 21, 2015. I am convinced that we will experience a fantastic conference, scientifically exciting and full of fond memories, in the unique environment of Lisbon.

Lluís Màrquez
EMNLP 2015 General Chair

Preface by the Program Committee Co-Chairs

August 25, 2015

Welcome to the 2015 Conference on Empirical Methods in Natural Language Processing! This year we received a record number of submissions. There were 1300 valid submissions. The 600 long papers and 700 short papers were allocated to one of 15 areas. The most popular areas this year were Semantics, Statistical Models and Machine Learning Methods, Text Mining and NLP applications, and Machine Translation.

Reviewing for a conference this size involves an enormous volunteer effort from many individuals. We are very grateful to our 30 area chairs and to the more than 900 researchers who reviewed the submissions. We accepted 312 papers (157 long and 155 short papers), representing a global acceptance rate of 24%. An additional 17 papers accepted by the TACL journal were presented at the conference as well.

To decide whether the accepted papers should be presented as talks or posters, we asked the area chairs, the reviewers, and the authors of accepted papers to vote on which papers they would like to attend. We showed the title of each paper and its abstract, but not its authors. 400 people provided their input. We selected talks based on popularity, while ensuring that each area was represented by at least one session. Our rationale for taking a vote was that papers that many people wanted to attend would be better served by presenting a talk in a large room, while papers with more specialized interest would benefit from the one-on-one interactions facilitated by posters. Rather than doing large plenary poster sessions, we have scheduled two parallel poster sessions with small batches of thematically similar papers that will be run simultaneously with the talks.

We selected best papers from a shortlist of 20 papers that were nominated by the area chairs. The best paper committee ranked the nominees, and based on their rankings we selected the following papers for the best paper awards:

- Best paper - *Broad-coverage CCG Semantic Parsing with AMR* by Yoav Artzi, Kenton Lee and Luke Zettlemoyer.
- Best paper - *Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems* by Tsung-Hsien Wen, Milica Gasic, Nikola Mrkšić, Pei-Hao Su, David Vandyke and Steve Young.

IBM has provided a cash scholarship for us to award to the best student paper. This will go to Tsung-Hsien Wen, since he is currently a student. The following papers received an honorable mention for the best paper award:

- Honorable mention for best paper - *Traversing Knowledge Graphs in Vector Space* by Kelvin Guu, John Miller and Percy Liang.
- Honorable mention for best paper - *Building a shared world: mapping distributional to model-theoretic semantic spaces* by Aurélie Herbelot and Eva Maria Vecchi.
- Honorable mention for best paper - *Language Understanding for Text-based Games using Deep Reinforcement Learning* by Karthik Narasimhan, Tejas Kulkarni and Regina Barzilay.
- Honorable mention for best short paper - *Joint Lemmatization and Morphological Tagging with Lemming* by Thomas Müller, Ryan Cotterell, Alexander Fraser and Hinrich Schütze.

- Honorable mention for best short paper - *Semi-Supervised Bootstrapping of Relationship Extractors with Distributional Semantics* by David S. Batista, Bruno Martins and Mário J. Silva.

This year we created a new "Best data set or resource" award, since so much work in our community is driven by data. The paper that receiving this inaugural distinction is:

- Best data set or resource - *A large annotated corpus for learning natural language inference* by Samuel R. Bowman, Gabor Angeli, Christopher Potts and Christopher D. Manning.

With two honorable mentions:

- Notable data set or resource - *That's So Annoying!!!: A Lexical and Frame-Semantic Embedding Based Data Augmentation Approach to Automatic Categorization of Annoying Behaviors using #petpeeve Tweets* by William Yang Wang and Diyi Yang.
- Notable data set or resource - *Modeling Reportable Events as Turning Points in Narrative* by Jessica Ouyang and Kathy McKeown.

We decided to give more awards than in past years by recognizing papers with honorable mentions and by creating the new best data or resource award. Our goal was to recognize roughly the top 1% of all of the submissions to the conference with awards (recognizing approximately the top 5% of accepted papers). We are very grateful to our invited speakers Yoshua Bengio and Justin Grimmer.

Yoshua Bengio is professor of Computer Science and Operations Research at the Université de Montréal. He is the author of two books and more than 200 publications, the most cited being in the areas of deep learning, recurrent neural networks, probabilistic learning algorithms, natural language processing and manifold learning. He co-directs the Canadian Institute for Advanced Research's program on deep learning. He is on the board of NIPS. Professor Bengio's research into deep learning has had a dramatic impact on the field of NLP in the past few years, and has invigorated interest in AI through machine learning.

Justin Grimmer is an associate professor of Political Science at Stanford University. His research uses statistical methods to examine American politics. He is the author of two books on the topic "Representational Style in Congress: What Legislators Say and Why It Matters" and "The Impression of Influence: How Legislator Communication and Government Spending Cultivate a Personal Vote." His work has appeared in the American Political Science Review, American Journal of Political Science, Journal of Politics, Political Analysis, Proceedings of the National Academy of Sciences, Regulation and Governance, and Poetics. Professor Grimmer's research points to exciting new directions for computational social science and how the field of NLP can facilitate research in many areas.

We thank them in advance for coming to the conference and sharing their insights.

We would also like to thank our general chair Lluís Màrquez, André Martins and João Graça and colleagues for their excellent work with the local organization, and Yuval Marton and Daniele Pighin for doing an excellent job assembling these proceedings.

We thank SIGDAT for inviting us to serve as Program Co-Chairs of EMNLP 2015. We hope that the conference is an excellent one. Enjoy your stay in Lisbon!

Chris Callison-Burch and Jian Su
EMNLP 2015 Program Committee Co-Chairs

Organizing Committee

General Chair
Lluís Màrquez, Qatar Computing Research Institute

Program Co-Chairs
Chris Callison-Burch, University of Pennsylvania
Jian Su, Institute for Infocomm Research (I2R)

Workshops Co-Chairs
Zornitsa Kozareva, Yahoo! Labs
Jörg Tiedemann, University of Helsinki

Tutorial Co-Chairs
Maggie Li, Hong Kong Polytechnic University
Khalil Sima'an, University of Amsterdam

Publication Co-Chairs
Daniele Pighin, Google Inc.
Yuval Marton, Microsoft Corp.

Publicity Chair
Barbara Plank, University of Copenhagen

Sponsorship Team
João Graça, Unbabel Inc.
Hang Li, Huawei Technologies (ISC Representative for EMNLP)

Student Scholarships Co-Chairs
Fancisco Guzmán, Qatar Computing Research Institute
Lluís Padró, Technical University of Catalonia

SIGDAT Liaison
Noah Smith, University of Washington

Local Organizing Committee

Local Co-Chairs
André Martins, Priberam
João Graça, Unbabel Inc.

Local Publicity Chair
Isabel Trancoso, University of Lisbon

Handbook Chair
Fernando Batista, University Institute of Lisbon (ISCTE-IUL)

Website and App Chair
Bruno Martins, University of Lisbon

Student Volunteer Coordinator
Luísa Coheur, University of Lisbon

Local Arrangements Chair
Helena Moniz, Instituto de Engenharia de Sistemas e Computadores - I&D (INESC-ID)

Program Committee

Program Co-Chairs

Chris Callison-Burch, University of Pennsylvania
Jian Su, Institute for Infocomm Research (I2R)

Area Chairs

Computational Psycholinguistics
Vera Demberg, Saarland University

Discourse, Dialogue, and Pragmatics
Kentaro Inui, Tohoku University
Oliver Lemon, Heriot Watt University

Information Extraction
Heng Ji, Rensselaer Polytechnic Institute
Sebastian Riedel, University College London
Jing Jiang, Singapore Management University

Information Retrieval and Question Answering
Zhu Xiaoyan, Tsinghua University
Liu Ting, Harbin Institute of Technology

Language and Vision
Julia Hockenmaier, University of Illinois Urbana Champaign

Machine Translation and Multilinguality
Zhang Min, SooChow University
Lucia Specia, University of Sheffield
Marine Carpuat, University of Maryland

NLP for the Web and Computational Social Science
Alice Oh, KAIST
Brendan O'Connor, University of Massachusetts Amherst

Phonology, Morphology, and Segmentation
Greg Kondrak, University of Alberta
Xu Sun, Peking University

Semantics
Marco Baroni, University of Trento
Shiqi Zhao, Baidu
Percy Liang, Stanford University

Sentiment Analysis and Opinion Mining
Ku Lun-Wei, Institute of Information Science, Academia Sinica
He Yulan, Aston University

Spoken Language Processing
Mari Ostendorf, University of Washington

Statistical Models and Machine Learning Methods for NLP
Trevor Cohn, University of Melbourne
Jordan Boyd-Graber, University of Colorado

Summarization and Generation
Meg Mitchell, Microsoft Research
Furu Wei, Microsoft Research Asia

Tagging, Chunking, Syntax and Parsing
Shay Cohen, University of Edinburgh
Yue Zhang, Singapore University of Technology and Design
Yusuke Miyao, National Institute of Informatics, Japan

Text Mining and NLP applications
Sophia Ananiadou, University of Manchester

Best Paper Award Committee
Douglas Downey, Northwestern University
Kevin Knight, University of Southern California
Qun Liu, Dublin City University
Miles Osborne, Bloomberg
Mari Ostendorf, University of Washington
Xiaojun Wan, Peking University
Janyce Weibe, University of Pittsburgh

Primary Reviewers

Balamurali A R; Omri Abend; Erman Acar; Karteek Addanki; Stergos Afantenos; Željko Agić; Lars Ahrenberg; Vicent Alabau; Mohammed Alam; Nikolaos Aletras; Alexandre Allauzen; Mariana S. C. Almeida; Yasemin Altun; Bharat Ram Ambati; Hadi Amiri; Waleed Ammar; Sophia Ananiadou; Jacob Andreas; Nicholas Andrews; David Andrzejewski; Gabor Angeli; Roitberg Anna; Stephen Anthony; Marianna Apidianaki; Ioannis Arapakis; Yuki Arase; Mihael Arcan; Yoav Artzi; Isabelle Augenstein; Eleftherios Avramidis; AiTi Aw; Wilker Aziz; Amittai Axelrod;

Xiao Bai; JinYeong Bak; Timothy Baldwin; Kalika Bali; Miguel Ballesteros; David Bamman; Pratyush Banerjee; Mohit Bansal; Libby Barak; Marco Baroni; Regina Barzilay; Roberto Basili; Emanuele Bastianelli; Fernando Batista; Riza Theresa Batista-Navarro; Daniel Bauer; Daniel Beck; Beata Beigman Klebanov; Núria Bel; David Belanger; Yonatan Belinkov; Islam Beltagy; Emily M. Bender; Meriem Bendris; Fabrício Benevenuto; Kepa Bengoetxea; Jonathan Berant; Taylor Berg-Kirkpatrick; Raffaella Bernardi; Nicola Bertoldi; Steven Bethard; Klinton Bicknell; Lidong Bing; Alexandra Birch; Arianna Bisazza; Yonatan Bisk; Alan W Black; Frédéric Blain; Eduardo Blanco; Bernd Bohnet; Gemma Boleda; Francis Bond; Kalina Bontcheva; Georgeta Bordea; Elizabeth Boschee; Houda Bouamor; Gerlof Bouma; Samuel R. Bowman; Jordan Boyd-Graber; S.R.K. Branavan; Jonathan Brennan; Chris Brew; Ted Briscoe; Chris Brockett; William Bryce; Christian Buck; Razvan Bunescu; David Burkett; David Buttler; Antoine Bordes;

José G. C. de Souza; Aoife Cahill; Chris Callison-Burch; Berkant Barla Cambazoglu; Nicola Cancedda; Marie Candito; Hailong Cao; Cornelia Caragea; Marine Carpuat; Xavier Carreras; Ben Carterette; Francisco Casacuberta; Tommaso Caselli; Vittorio Castelli; Giuseppe Castellucci; Asli Celikyilmaz; Daniel Cer; Mauro Cettolo; Arun Chaganty; Nathanael Chambers; Yee Seng Chan; Baobao Chang; Jonathan Chang; Kai-Wei Chang; Ming-Wei Chang; Yin-Wen Chang; Wanxiang Che; Bin Chen; Boxing Chen; Chen Chen; Danqi Chen; Hsin-Hsi Chen; Liwei Chen; Wenliang Chen; Yun-Nung Chen; Zhiyuan Chen; Colin Cherry; Sean Chester; Jackie Chi Kit Cheung; David Chiang; Christian Chiarcos; Hai Leong Chieu; Do Kook Choe; Eunsol Choi; Jinho D. Choi; Yejin Choi; Monojit Choudhury; Christos Christodoulopoulos; Grzegorz Chrupała; Chenhui Chu; Tagyoung Chung; Mark Cieliebak; Stephen Clark; Ann Clifton; Martin Cmejrek; Moreno Coco; Shay B. Cohen; Trevor Cohn; Paul Cook; Bonaventura Coppola; Greg Coppola; Marta R. Costa-Jussà; Benoit Crabbé; Danilo Croce; Fabien Cromieres; James Cross; Montse Cuadros; Heriberto Cuayahuitl; Lei Cui; Ronan Cummins;

Zhicheng Dou; Walter Daelemans; Daniel Dahlmeier; Van Dang; William Darling; Dipanjan Das; Pradipto Das; Rajarshi Das; Sajib Dasgupta; Pradeep Dasigi; Steve DeNeefe; John DeNero; Judith Degen; Felice Dell'Orletta; Vera Demberg; Dina Demner-Fushman; Pascal Denis; Michael Denkowski; Tejaswini Deoskar; Leon Derczynski; Aliya Deri; Nina Dethlefs; Laura Dietz; Georgiana Dinu; Qing Dou; Doug Downey; Eduard Dragut; Markus Dreyer; Gregory Druck; Lan Du; Xiangyu Duan; Kevin Duh; Nadir Durrani; Greg Durrett; Chris Dyer; Marc Dymetman; Bharath Dandala; Adrià de Gispert; Gerard de Melo; Ann Devitt;

Richard Eckart de Castilho; Patrick Ehlen; Carsten Eickhoff; Andreas Eisele; Jacob Eisenstein; Ismail El Maarouf; Michael Elhadad; Micha Elsner; Tomaž Erjavec; Arash Eshghi; Cristina España-Bonet; Luis Espinosa Anke; Miquel Esplà-Gomis;

Licheng Fang; Benamara Farah; Noura Farra; Manaal Faruqui; Benoit Favre; Anna Feldman; Minwei Feng; Song Feng; Yang Feng; Yansong Feng; Raquel Fernandez; Daniel Fernández-González; Francis Ferraro; Michele Filannino; Simone Filice; Denis Filimonov; Nicholas FitzGerald; Darja Fišer; Jeffrey

Flanigan; Tiziano Flati; Radu Florian; John Foley; José A. R. Fonollosa; Mikel Forcada; Lluís Formiga; Eric Fosler-Lussier; George Foster; Stefan L. Frank; Stella Frank; Markus Freitag; Lea Frermann; Annemarie Friedrich; Guohong Fu; Hagen Fuerstenau; Alona Fyshe;

Hongyu Guo; Michel Galley; Matthias Gallé; Michael Gamon; Kavita Ganesan; Wei Gao; Claire Gardent; Dan Garrette; Milica Gasic; Jon Gauthier; Tao Ge; Spandana Gella; Kallirroi Georgila; Alborz Geramifard; Debanjan Ghosh; Daniel Gildea; Kevin Gimpel; Filip Ginter; Roxana Girju; Koldo Gojenola; Yoav Goldberg; Dan Goldwasser; Zhengxian Gong; Graciela Gonzalez; Jesús González-Rubio; Michael Wayne Goodman; Amit Goyal; Natalia Grabar; Mark Granroth-Wilding; Spence Green; Stephan Greene; Edward Grefenstette; Cyril Grouin; Stig-Arne Grönroos; Fritz Guenther; Jiafeng Guo; Jiang Guo; Weiwei Guo; Sonal Gupta; Iryna Gurevych; Jose Manuel Gómez; Carlos Gómez-Rodríguez;

Ben Hachey; Barry Haddow; Gholamreza Haffari; Masato Hagiwara; Dilek Hakkani-Tur; David Hall; Keith Hall; Thierry Hamon; Bo Han; Shuguang Han; Xianpei Han; Greg Hanneman; Yu Hao; Sanda Harabagiu; Christian Hardmeier; Saša Hasan; Chikara Hashimoto; Kazuma Hashimoto; Eva Hasler; Jun Hatori; He He; Xiaodong He; Yifan He; Yulan He; Kenneth Heafield; Larry Heck; Carmen Heger; Michael Heilman; James Henderson; John Henderson; Aurélie Herbelot; Karl Moritz Hermann; Felix Hieber; Derrick Higgins; Gerold Hintz; Graeme Hirst; Hieu Hoang; Julia Hockenmaier; Yu Hong; Ales Horak; Julian Hough; Dirk Hovy; Chun-Nan Hsu; Yuening Hu; Fei Huang; Hongzhao Huang; Liang Huang; Minlie Huang; Ruihong Huang; Xuanjing Huang; Yi-Ting Huang; Zhongqiang Huang; Matthias Huck; Rebecca Hwa;

Gonzalo Iglesias; Ryu Iida; Shajith Ikbal; Iustina Ilisei; Diana Inkpen; Naoya Inoue; Kentaro Inui; Aminul Islam;

Guillaume Jacquet; Line Jakubiec-Jamet; Myungha Jang; Peter Jansen; Sujay Kumar Jauhar; Arun Kumar Jayapal; Laura Jehl; Heng Ji; Yangfeng Ji; Jing Jiang; Wenbin Jiang; Peng Jin; Anders Johannsen; Richard Johansson; Melvin Johnson Premkumar; Dan Jurafsky; David Jurgens;

Mijail Kabadjov; Nobuhiro Kaji; Nal Kalchbrenner; Justine Kao; Dimitri Kartsaklis; Saurabh Kataria; Daisuke Kawahara; Anna Kazantseva; Simon Keizer; Frank Keller; Casey Kennington; Maxim Khalilov; Ankur Khanna; Mitesh M. Khapra; Douwe Kiela; Halil Kilicoglu; Dongho Kim; Dongwoo Kim; Jin-Dong Kim; Seungyeon Kim; Suin Kim; Chunyu Kit; Alexandre Klementiev; Roman Klinger; Ari Kobren; Łukasz Kobyliński; Philipp Koehn; Rob Koeling; Mamoru Komachi; Grzegorz Kondrak; Fang Kong; Lingpeng Kong; Ioannis Konstas; Georgios Kontonatsios; Ioannis Korkontzelos; Zornitsa Kozareva; Mikhail Kozhevnikov; Martin Krallinger; Jayant Krishnamurthy; Kriste Krstovski; Canasai Kruengkrai; Germán Kruszewski; Lun-Wei Ku; Roland Kuhn; Anjishnu Kumar; Shankar Kumar; Jonathan K. Kummerfeld; Tsung-Ting Kuo; Nate Kushman; Tom Kwiatkowski;

Mathias Lambert; Patrik Lambert; Vasileios Lampos; Gerasimos Lampouras; Man Lan; Yanyan Lan; Jun Lang; Phillippe Langlais; Mirella Lapata; First Name Last Name; Jey Han Lau; Angeliki Lazaridou; Svetlana Lazebnik; Dieu-Thu Le; Joseph Le Roux; Kenton Lee; Lung-Hao Lee; Sungjin Lee; Tao Lei; Oliver Lemon; Alessandro Lenci; Chee Wee Leong; Gina-Anne Levow; Omer Levy; Mike Lewis; Binyang Li; Chen Li; Chenliang Li; Fangtao Li; Haibo Li; Hao Li; Jing Li; Jiwei Li; Junhui Li; Peifeng Li; Qi Li; Shoushan Li; Sujian Li; Wenjie Li; Yanen Li; Zhenghua Li; Maria Liakata; Percy Liang; Zhen Liao; Chenghua Lin; Chu-Cheng Lin; Xi Victoria Lin; Wang Ling; Xiao Ling; Tal Linzen; Bing Liu; Fei Liu; Jialu Liu; Jing Liu; Kang Liu; Qun Liu; Rui Liu; Ting Liu; Xiaodong Liu; Yang Liu; Yang Liu; Yiqun Liu; Zhanyi Liu; Zhiyuan Liu; Nikola Ljubešić; Chi-kiu Lo; Pavel Logacev; Varvara Logacheva; Oier Lopez de Lacalle; Bin Lu; Wei Lu; Xiaofei Lu; Marco Lui; Zhunchen Luo; Thang Luong;

Véronique Moriceau; Ji Ma; Yanjun Ma; Zongyang Ma; Wolfgang Macherey; Pranava Swaroop Madhyastha; Nitin Madnani; Pierre Magistry; Gideon Mann; Christopher D. Manning; Amin Mantrach; Xian-Ling Mao; Georgiana Marsic; Scott Martin; André F. T. Martins; Yuval Marton; Héctor Martínez Alonso; Prashant Mathur; Takuya Matsuzaki; Arne Mauser; Jonathan May; David McClosky; John Philip McCrae; Christopher McCubbin; Ryan McDonald; Louise McNally; Avihai Mejer; Yelena Mejova; Oren Melamud; Pablo Mendes; Florian Metze; Haitao Mi; Rada Mihalcea; Claudiu Mihăilă; David Mimno; Bonan Min; Zhao-Yan Ming; Shachar Mirkin; Seyed Abolghasem Mirroshandel; Paramita Mirza; Margaret Mitchell; Makoto Miwa; Yusuke Miyao; Daichi Mochihashi; Saif Mohammad; Mitra Mohtarami; Karo Moilanen; Christian Monson; Manuel Montes; Christof Monz; Taesun Moon; Raymond Mooney; Roser Morante; Alessandro Moschitti; Arjun Mukherjee; Philippe Muller; Dragos Munteanu; Yugo Murawaki; Brian Murphy; Lluís Màrquez;

Jian-Yun Nie; Maria Nadejde; Vinita Nahar; Preslav Nakov; Karthik Narasimhan; Shashi Narayan; Tahira Naseem; Vivi Nastase; Borja Navarro; Roberto Navigli; Raheel Nawaz; Mark-Jan Nederhof; Arvind Neelakantan; Sapna Negi; Aida Nematzadeh; Graham Neubig; Hwee Tou Ng; Vincent Ng; Nhung Nguyen; Viet-An Nguyen; Garrett Nicolai; Jan Niehues; Rodney Nielsen; Zheng-Yu Niu; Joakim Nivre; Joel Nothman; Scott Nowson; Pierre Nugues; Malte Nuhn;

Brendan O'Connor; Timothy O'Donnell; Stephan Oepen; Kemal Oflazer; Alice Oh; Naoaki Okazaki; Desmond Ong; Constantin Orasan; Vicente Ordonez; Daniel Ortiz-Martínez; Miles Osborne; Mari Ostendorf;

Deepak P; Ulrike Pado; Muntsa Padró; Gustavo Paetzold; Georgios Paliouras; Martha Palmer; Alessio Palmero Aprosio; Sinno J. Pan; Denis Paperno; Aasish Pappu; Rebecca J. Passonneau; Panupong Pasupat; Siddharth Patwardhan; Michael J. Paul; Adam Pauls; Ellie Pavlick; Lisa Pearl; Stephan Peitz; Filipa Peleja; Xiaochang Peng; Sergio Penkale; Gerald Penn; Bianca Pereira; Bryan Perozzi; Johann Petrak; Slav Petrov; Anselmo Peñas; Nghia The Pham; Lawrence Phillips; Karl Pichotta; Daniele Pighin; Mohammad Taher Pilehvar; Manfred Pinkal; Juan Pino; Yuval Pinter; Emily Pitler; Barbara Plank; Massimo Poesio; Tamara Polajnar; Heather Pon-Barry; Simone Paolo Ponzetto; Andrei Popescu-Belis; Maja Popović; Vinodkumar Prabhakaran; John Prager; Rashmi Prasad; Daniel Preoţiuc-Pietro; Prokopis Prokopidis; Matthew Purver; Verónica Pérez-Rosas;

Longhua Qian; Xian Qian; Lu Qin; Likun Qiu; Long Qiu; Xipeng Qiu; Lizhen Qu; Chris Quirk;

Ella Rabinovich; Will Radford; Preethi Raghavan; Altaf Rahman; Rafal Rak; Carlos Ramisch; Mohammad Sadegh Rasooli; Siva Reddy; Sravana Reddy; Marek Rei; Roi Reichart; Nils Reiter; Xiang Ren; Philip Resnik; Corentin Ribeyre; Kyle Richardson; Sebastian Riedel; Martin Riedl; Stefan Riezler; German Rigau; Laura Rimell; Fabio Rinaldi; Miguel Rios; Alan Ritter; Brian Roark; Angus Roberts; Kirk Roberts; Tim Rocktäschel; Horacio Rodriguez; Marcus Rohrbach; Stephen Roller; Francesco Ronzano; Andrew Rosenberg; Paolo Rosso; Michael Roth; Johann Roturier; Mickael Rouvier; Alla Rozovskaya; Eugen Ruppert;

Jun Sun; Markus Saers; Saurav Sahay; Hassan Saif; Hassan Sajjad; Keisuke Sakaguchi; Mohammad Salameh; Diarmuid Ó Séaghdha; Bahar Salehi; Tanja Samardzic; Mark Sammons; Germán Sanchis Trilles; Federico Sangati; Baskaran Sankaran; Motoki Sano; Felix Sasaki; Ryohei Sasano; Asad Sayeed; Carolina Scarton; David Schlangen; Natalie Schluter; Helmut Schmid; Nathan Schneider; Michael Schuhmacher; William Schuler; H. Andrew Schwartz; Lane Schwartz; Roy Schwartz; Holger Schwenk; Hinrich Schütze; Djamé Seddah; Satoshi Sekine; Jean Senellart; Rico Sennrich; Hendra Setiawan; Burr

Settles; Aliaksei Severyn; Serge Sharoff; Shuming Shi; Chaitanya Shivade; Maryam Siahbani; Carina Silberer; Yanchuan Sim; Khalil Sima'an; Patrick Simianer; Kiril Simov; Sameer Singh; Kevin Small; Peter Smit; Noah A. Smith; Jan Šnajder; Artem Sokolov; Swapna Somasundaran; Linfeng Song; Lucia Specia; Vivek Srikumar; Miloš Stanojević; Mark Steedman; Pontus Stenetorp; Brandon Stewart; Veselin Stoyanov; Carlo Strapparava; Karl Stratos; Emma Strubell; Jannik Strötgen; Jian Su; Qi Su; Kazunari Sugiyama; Le Sun; Weiwei Sun; Xu Sun; Mihai Surdeanu; Hisami Suzuki; Swabha Swayamdipta; György Szarvas; Stan Szpakowicz; Idan Szpektor; Felipe Sánchez-Martínez; Anders Søgaard; Patrick Saint-Dizier; Jinsong Su;

Hiroya Takamura; Partha P. Talukdar; Duyu Tang; Jian Tang; Jiliang Tang; Xavier Tannier; Kapil Thadani; Sam Thomson; Jörg Tiedemann; Christoph Tillmann; Ivan Titov; Nadi Tomeh; Sara Tonelli; Fatemeh Torabi Asr; Antonio Toral; Kristina Toutanova; Jun'ichi Tsujii; Oren Tsur; Yoshimasa Tsuruoka; Yulia Tsvetkov; Gokhan Tur; Francis Tyers; Oscar Täckström;

Raghavendra Udupa; Lyle Ungar; Olga Uryupina; Jakob Uszkoreit;

Sowmya Vajjala; Benjamin Van Durme; Daniele Vannella; Ashish Vaswani; Marc Verhagen; Yannick Versley; Laure Vieu; David Vilar; Luke Vilnis; Sami Virpioja; Karthik Visweswariah; Andreas Vlachos; Rob Voigt; Svitlana Volkova; Ivan Vulić; Vinod Vydiswaran; Marieke van Erp; Marten van Schijndel; Antal van den Bosch;

Houfeng Wang; Joachim Wagner; Aurelien Waite; Xiaojun Wan; Chi Wang; Chuan Wang; Hongling Wang; Hongning Wang; Lu Wang; Rui Wang; Sida I. Wang; Wen Wang; William Yang Wang; Zhiguo Wang; Zhongqing Wang; Zhuoran Wang; Taro Watanabe; Aleksander Wawer; Kellie Webster; Furu Wei; Zhongyu Wei; Daniel Weld; Janyce Wiebe; Michael Wiegand; John Wieting; Jason D Williams; Steven Wilson; Shuly Wintner; Guillaume Wisniewski; Silke Witt-Ehsani; Travis Wolfe; Kam-Fai Wong; Alina Wróblewska; Stephen Wu; Xianchao Wu; Yuanbin Wu; Joern Wuebker;

Yunqing Xia; Tong Xiao; Boyi Xie; Pengtao Xie; Deyi Xiong; Jia Xu; Liheng Xu; Ruifeng Xu; Wei Xu; Wenduan Xu; Ying Xu;

Rui Yan; Bin Yang; Weiwei Yang; Yaqin Yang; Yi Yang; Yi Yang; Helen Yannakoudakis; Jin-ge Yao; Xuchen Yao; Ainur Yessenalina; Xing Yi; Wen-tau Yih; Seid Muhie Yimam; Dani Yogatama; Bei Yu; Dianhai Yu; Liang-Chih Yu; Mo Yu; Zhou Yu; François Yvon; Tae Yano;

Deyu Zhou; Manzil Zaheer; Omar Zaidan; Fabio Massimo Zanzotto; Alessandra Zarcone; Sina Zarrieß; Rabih Zbib; Richard Zens; Torsten Zesch; Luke Zettlemoyer; Feifei Zhai; Ke Zhai; Jiajun Zhang; Kevin Zhang; Meishan Zhang; Min Zhang; Min Zhang; Qi Zhang; Wei Zhang; Wei-Nan Zhang; Yu Zhang; Yuan Zhang; Yuchen Zhang; Yue Zhang; Zhe Zhang; Zhichang Zhang; Hai Zhao; Jun Zhao; Kai Zhao; Qiuye Zhao; Shiqi Zhao; Wayne Xin Zhao; Xin Zhao; Yanyan Zhao; Zhi Zhong; Guangyou Zhou; Guodong Zhou; Junsheng Zhou; Mianwei Zhou; Xinjie Zhou; Jingbo Zhu; Muhua Zhu; Xiaodan Zhu; Imed Zitouni; Bowei Zou; Arkaitz Zubiaga; Pierre Zweigenbaum; Xiaoyan Zhu;

Additional Reviewers

Abeed Sarker; Abouenour Lahsen; Afra Alishahi; Afshin Rahimi; Aitor García Pablos; Aitziber Atutxa Salazar; Alan Lee; Alicia Pérez Ramírez; Anders Johannsen; Anders Søgaard; Andreas Grivas; Anna Jørgensen; Annemarie Friedrich; Antonio Uva; Austin Matthews; Avihai Mejer; Baolin Peng; Benjamin Marie; Benjamin Yap; Bo Han; Bonnie Webber; Bradley Hauer; Camilo Thorne; Chong Zhou; Daniel Torregrosa; David Buttler; Davy Weissenbacher; Desmond Ong; Ehsan Emazadeh; Elias Zavitsanos; Enrique Vidal; Fandong Meng; Federico Nanni; Filipa Peleja; Francis Ferraro; Francisco Rangel; Gözde Özbal; Henry Anaya-Sanchez; Hieu Pham; Ignacio Iacobacci; I-Ta Lee; Jack Hessel; James Cross; Jing Li; Jing Ma; Jintao Ye; Joost Bastings; Kai Zhao; Kateryna Tymoshenko; Kazuya Kawakami; Ke Tao; Ke Zhai; Keith Stevens; Lane Schwartz; Lasse Borgholt; Leila Wehbe; Long Duong; Lucie Flekova; Marc Franco-Salvador; Marco Del Tredici; Maximin Coavoux; Meng Zhang; Michal Lukasik; Mingxuan Wang; Mohamed Mouine; Mohammad Salameh; Muhua Zhu; Nadir Durrani; Nan Yang; Nichola Lubold; Nima Pourdamghani; Nina Zhou; Phil Williams; Prodromos Malakasiotis; Rachel Rudinger; Rui Xia; Serra Sinem Tekiroglu; Shachi H Kumar; Shruti Rijhwani; Songxian Xie; Stas Semeniuta; Stefano Faralli; Stephan Gouws; Stephen Maddock; Sungjoon Park; Susana Palmaz; Tobias Domhan; Tomer Levinboim; Vaggelis Michelioudakis; Vassilis Papavassiliou; Vivek Kulkarni; Wang Ling; Wen Wang; Weston Feely; Wu Kui; Xanda Schofield; Xiang Li; Xiaodong He; Yan Fu; Yanran Li; Yi Luan; Yijia Liu; Yulia Tsvetkov; Zhihua Zhang.

Invited Speaker: Yoshua Bengio

Deep Learning of Semantic Representations

Abstract: The core ingredient of deep learning is the notion of distributed representation. This talk will start by explaining its theoretical advantages, in comparison with non-parametric methods based on counting frequencies of occurrence of observed tuples of values (like with n-grams). The talk will then explain how having multiple levels of representation, i.e., depth, can in principle give another exponential advantage. Neural language models have been extremely successful in recent years but extending their reach from language modeling to machine translation is very appealing because it forces the learned intermediate representations to capture meaning, and we found that the resulting word embeddings are qualitatively different. Recently, we introduced the notion of attention-based encoder-decoder systems, with impressive results on machine translation several language pairs and for mapping an image to a sentence, and these results will conclude the talk.

Biography: Yoshua Bengio received a PhD in Computer Science from McGill University, Canada in 1991. After two post-doctoral years, one at M.I.T. with Michael Jordan and one at AT&T Bell Laboratories with Yann LeCun and Vladimir Vapnik, he became professor at the Department of Computer Science and Operations Research at Université de Montréal. He is the author of two books and more than 200 publications, the most cited being in the areas of deep learning, recurrent neural networks, probabilistic learning algorithms, natural language processing and manifold learning. He is among the most cited Canadian computer scientists and is or has been associate editor of the top journals in machine learning and neural networks. Since '2000 he holds a Canada Research Chair in Statistical Learning Algorithms, since '2006 an NSERC Industrial Chair, since '2005 his is a Senior Fellow of the Canadian Institute for Advanced Research and since 2014 he co-directs its program focused on deep learning. He is on the board of the NIPS foundation and has been program chair and general chair for NIPS. He has co-organized the Learning Workshop for 14 years and co-created the new International Conference on Learning Representations. His current interests are centered around a quest for AI through machine learning, and include fundamental questions on deep learning and representation learning, the geometry of generalization in high-dimensional spaces, manifold learning, biologically inspired learning algorithms, and challenging applications of statistical machine learning.

Invited Speaker: Justin Grimmer

Measuring How Elected Officials and Constituents Communicate

Abstract: This talk will show how elected officials use communication to cultivate support with constituents, how constituents express their views to elected officials, and why biases in both kinds of communication matter for political representation. To demonstrate the bias and its effects, I propose to use novel collections of political texts and new text as data methods. Using the new data and methods, I will show how the incentives of communication contribute to perceptions of an angry public and vitriolic politicians. Among elected officials, the ideologically extreme members of Congress disproportionately participate in policy debates, resulting in political debates that occur between the most extreme members of each party. Among constituents, the most ideologically extreme and angry voters disproportionately contact their member of Congress, creating the impression of a polarized and vitriolic public. The talk will explain how the findings help us to understand how representation occurs in American politics, while also explaining how computational tools can help address questions in the social sciences.

Biography: Justin Grimmer is an associate professor of political science at Stanford University. His research examines how representation occurs in American politics using new statistical methods. His first book Representational Style in Congress: What Legislators Say and Why It Matters (Cambridge University Press, 2013) shows how senators define the type of representation they provide constituents and how this affects constituents' evaluations and was awarded the 2014 Richard Fenno Prize. His second book The Impression of Influence: How Legislator Communication and Government Spending Cultivate a Personal Vote (Princeton University Press, 2014 with Sean J. Westwood and Solomon Messing) demonstrates how legislators ensure they receive credit for government actions. His work has appeared in the American Political Science Review, American Journal of Political Science, Journal of Politics, Political Analysis, Proceedings of the National Academy of Sciences, Regulation and Governance, and Poetics.

Conference Program

Friday, September 18, 2015

18:30–20:00 Welcome Reception

Saturday, September 19, 2015

07:30–18:00 Registration

08:40–10:00 Session P1: Plenary Session

08:40–09:00 *Opening Remarks and Introductory Speeches*
General Chair, Program Co-Chairs and Local Co-Chairs

09:00–10:00 *Invited Talk: Deep Learning of Semantic Representations*
Yoshua Bengio

10:00–10:30 *Coffee break*

10:30–12:10 Session 1A: Semantics (Long + TACL Papers)

10:30–10:55 *Language Understanding for Text-based Games using Deep Reinforcement Learning*
Karthik Narasimhan, Tejas Kulkarni and Regina Barzilay

10:55–11:20 *Distributional vectors encode referential attributes*
Abhijeet Gupta, Gemma Boleda, Marco Baroni and Sebastian Padó

11:20–11:45 *Building a shared world: mapping distributional to model-theoretic semantic spaces*
Aurélie Herbelot and Eva Maria Vecchi

11:45–12:10 *[TACL] Deriving Boolean Structures from Distributional Vectors*
Germán Kruszewski, Denis Paperno and Marco Baroni

Saturday, September 19, 2015 (continued)

10:30–12:10 Session 1B: Machine Translation (Long + TACL Papers)

10:30–10:55 *Dependency Graph-to-String Translation*
Liangyou Li, Andy Way and Qun Liu

10:55–11:20 *Reordering Grammar Induction*
Miloš Stanojević and Khalil Sima'an

11:20–11:45 *Syntax-based Rewriting for Simultaneous Machine Translation*
He He, Alvin Grissom II, John Morgan, Jordan Boyd-Graber and Hal Daumé III

11:45–12:10 *[TACL] Modelling and Optimizing on Syntactic N-Grams for Statistical Machine Translation*
Rico Sennrich

10:30–12:10 Session 1C: NLP for Web and Social Media, including Computational Social Science (Long Papers)

10:30–10:55 *Identifying Political Sentiment between Nation States with Social Media*
Nathanael Chambers, Victor Bowen, Ethan Genco, Xisen Tian, Eric Young, Ganesh Harihara and Eugene Yang

10:55–11:20 *Open Extraction of Fine-Grained Political Statements*
David Bamman and Noah A. Smith

11:20–11:45 *Using Personal Traits For Brand Preference Prediction*
Chao Yang, Shimei Pan, Jalal Mahmud, Huahai Yang and Padmini Srinivasan

11:45–12:10 *Semantic Annotation for Microblog Topics Using Wikipedia Temporal Information*
Tuan Tran, Nam Khanh Tran, Asmelash Teka Hadgu and Robert Jäschke

10:30–12:10 Session 1D (P1-4): Summarization (Long Paper Posters)

System Combination for Multi-document Summarization
Kai Hong, Mitchell Marcus and Ani Nenkova

Phrase-based Compressive Cross-Language Summarization
Jin-ge Yao, Xiaojun Wan and Jianguo Xiao

Re-evaluating Automatic Summarization with BLEU and 192 Shades of ROUGE
Yvette Graham

Indicative Tweet Generation: An Extractive Summarization Problem?
Priya Sidhaye and Jackie Chi Kit Cheung

Saturday, September 19, 2015 (continued)

10:30–12:10 Session 1D (P5): Language and Vision (Long Paper Posters)

Visual Bilingual Lexicon Induction with Transferred ConvNet Features
Douwe Kiela, Ivan Vulić and Stephen Clark

10:30–12:10 Session 1D (P6-9): Sentiment Analysis and Opinion Mining (Long Paper Posters)

Cross Lingual Sentiment Analysis using Modified BRAE
Sarthak Jain and Shashank Batra

Monotone Submodularity in Opinion Summaries
Jayanth Jayanth, Jayaprakash Sundararaj and Pushpak Bhattacharyya

Joint Prediction for Entity/Event-Level Sentiment Analysis using Probabilistic Soft Logic Models
Lingjia Deng and Janyce Wiebe

Learning to Recognize Affective Polarity in Similes
Ashequl Qadir, Ellen Riloff and Marilyn Walker

10:30–12:10 Session 1E (P1-3): Language and Vision (Short Paper Posters)

Cross-document Event Coreference Resolution based on Cross-media Features
Tongtao Zhang, Hongzhi Li, Heng Ji and Shih-Fu Chang

A Survey of Current Datasets for Vision and Language Research
Francis Ferraro, Nasrin Mostafazadeh, Ting-Hao Huang, Lucy Vanderwende, Jacob Devlin, Michel Galley and Margaret Mitchell

Combining Geometric, Textual and Visual Features for Predicting Prepositions in Image Descriptions
Arnau Ramisa, Josiah Wang, Ying Lu, Emmanuel Dellandrea, Francesc Moreno-Noguer and Robert Gaizauskas

10:30–12:10 Session 1E (P4-16): Statistical Models and Machine Learning Methods (Short Paper Posters)

On A Strictly Convex IBM Model 1
Andrei Simion, Michael Collins and Cliff Stein

Factorization of Latent Variables in Distributional Semantic Models
Arvid Österlund, David Ödling and Magnus Sahlgren

Non-lexical neural architecture for fine-grained POS Tagging
Matthieu Labeau, Kevin Löser and Alexandre Allauzen

Online Representation Learning in Recurrent Neural Language Models
Marek Rei

A Model of Zero-Shot Learning of Spoken Language Understanding
Majid Yazdani and James Henderson

Modeling Tweet Arrival Times using Log-Gaussian Cox Processes
Michal Lukasik, P. K. Srijith, Trevor Cohn and Kalina Bontcheva

Pre-Computable Multi-Layer Neural Network Language Models
Jacob Devlin, Chris Quirk and Arul Menezes

Birds of a Feather Linked Together: A Discriminative Topic Model using Link-based Priors
Weiwei Yang, Jordan Boyd-Graber and Philip Resnik

Aligning Knowledge and Text Embeddings by Entity Descriptions
Huaping Zhong, Jianwen Zhang, Zhen Wang, Hai Wan and Zheng Chen

An Empirical Analysis of Optimization for Max-Margin NLP
Jonathan K. Kummerfeld, Taylor Berg-Kirkpatrick and Dan Klein

Learning Better Embeddings for Rare Words Using Distributional Representations
Irina Sergienya and Hinrich Schütze

Composing Relationships with Translations
Alberto Garcia-Duran, Antoine Bordes and Nicolas Usunier

Noise or additional information? Leveraging crowdsource annotation item agreement for natural language tasks.
Emily Jamison and Iryna Gurevych

12:10–13:30 ***Lunch***

Saturday, September 19, 2015 (continued)

13:30–15:10 Session 2A: Statistical Models and Machine Learning Methods (Long + TACL Papers)

13:30–13:55 *Evaluation methods for unsupervised word embeddings*
Tobias Schnabel, Igor Labutov, David Mimno and Thorsten Joachims

13:55–14:20 *Efficient Methods for Incorporating Knowledge into Topic Models*
Yi Yang, Doug Downey and Jordan Boyd-Graber

14:20–14:45 *Traversing Knowledge Graphs in Vector Space*
Kelvin Guu, John Miller and Percy Liang

14:45–15:10 *[TACL] Improving Topic Models with Latent Feature Word Representations*
Dat Quoc Nguyen, Richard Billingsley, Lan Du and Mark Johnson

13:30–15:10 Session 2B: Tagging, Chunking and Parsing (Long +TACL Papers)

13:30–13:55 *Density-Driven Cross-Lingual Transfer of Dependency Parsers*
Mohammad Sadegh Rasooli and Michael Collins

13:55–14:20 *A Neural Network Model for Low-Resource Universal Dependency Parsing*
Long Duong, Trevor Cohn, Steven Bird and Paul Cook

14:20–14:45 *Improved Transition-based Parsing by Modeling Characters instead of Words with LSTMs*
Miguel Ballesteros, Chris Dyer and Noah A. Smith

14:45–15:10 *[TACL] Approximation-Aware Dependency Parsing by Belief Propagation*
Matthew R. Gormley, Mark Dredze and Jason Eisner

13:30–15:10 Session 2C: Summarization (Long Papers)

13:30–13:55 *Sentence Compression by Deletion with LSTMs*
Katja Filippova, Enrique Alfonseca, Carlos A. Colmenares, Lukasz Kaiser and Oriol Vinyals

13:55–14:20 *An Empirical Comparison Between N-gram and Syntactic Language Models for Word Ordering*
Jiangming Liu and Yue Zhang

14:20–14:45 *A Neural Attention Model for Abstractive Sentence Summarization*
Alexander M. Rush, Sumit Chopra and Jason Weston

14:45–15:10 *Scientific Article Summarization Using Citation-Context and Article's Discourse Structure*
Arman Cohan and Nazli Goharian

13:30–15:10 Session 2D (P1-9): Text Mining and NLP Applications (Long Paper Posters)

Hashtag Recommendation Using Dirichlet Process Mixture Models Incorporating Types of Hashtags
Yeyun Gong, Qi Zhang and Xuanjing Huang

A Graph-based Readability Assessment Method using Word Coupling
Zhiwei Jiang, Gang Sun, Qing Gu, Tao Bai and Daoxu Chen

More Features Are Not Always Better: Evaluating Generalizing Models in Incident Type Classification of Tweets
Axel Schulz, Christian Guckelsberger and Benedikt Schmidt

Flexible Domain Adaptation for Automated Essay Scoring Using Correlated Linear Regression
Peter Phandi, Kian Ming A. Chai and Hwee Tou Ng

Show Me Your Evidence - an Automatic Method for Context Dependent Evidence Detection
Ruty Rinott, Lena Dankin, Carlos Alzate Perez, Mitesh M. Khapra, Ehud Aharoni and Noam Slonim

Spelling Correction of User Search Queries through Statistical Machine Translation
Saša Hasan, Carmen Heger and Saab Mansour

Human Evaluation of Grammatical Error Correction Systems
Roman Grundkiewicz, Marcin Junczys-Dowmunt and Edward Gillian

Learning a Deep Hybrid Model for Semi-Supervised Text Classification
Alexander Ororbia II, C. Lee Giles and David Reitter

Joint Embedding of Query and Ad by Leveraging Implicit Feedback
Sungjin Lee and Yifan Hu

Saturday, September 19, 2015 (continued)

13:30–15:10 Session 2E (P1-11): Information Extraction (Short Paper Posters)

Automatic Extraction of Time Expressions Accross Domains in French Narratives
Mike Donald Tapi Nzali, Xavier Tannier and Aurelie Neveol

Semi-Supervised Bootstrapping of Relationship Extractors with Distributional Semantics
David S. Batista, Bruno Martins and Mário J. Silva

Extraction and generalisation of variables from scientific publications
Erwin Marsi and Pinar Öztürk

Named entity recognition with document-specific KB tag gazetteers
Will Radford, Xavier Carreras and James Henderson

"A Spousal Relation Begins with a Deletion of engage and Ends with an Addition of divorce": Learning State Changing Verbs from Wikipedia Revision History
Derry Tanti Wijaya, Ndapandula Nakashole and Tom Mitchell

Improving Distant Supervision for Information Extraction Using Label Propagation Through Lists
Lidong Bing, Sneha Chaudhari, Richard Wang and William Cohen

An Entity-centric Approach for Overcoming Knowledge Graph Sparsity
Manjunath Hegde and Partha P. Talukdar

Semantic Relation Classification via Convolutional Neural Networks with Simple Negative Sampling
Kun Xu, Yansong Feng, Songfang Huang and Dongyan Zhao

A Baseline Temporal Tagger for all Languages
Jannik Strötgen and Michael Gertz

Named Entity Recognition for Chinese Social Media with Jointly Trained Embeddings
Nanyun Peng and Mark Dredze

Inferring Binary Relation Schemas for Open Information Extraction
Kangqi Luo, Xusheng Luo and Kenny Zhu

13:30–15:10 **Session 2E (P12-16): Information Retrieval and Question Answering (Short Paper Posters)**

LDTM: A Latent Document Type Model for Cumulative Citation Recommendation
Jingang Wang, Dandan Song, Zhiwei Zhang, Lejian Liao, Luo Si and Chin-Yew Lin

Online Sentence Novelty Scoring for Topical Document Streams
Sungjin Lee

Global Thread-level Inference for Comment Classification in Community Question Answering
Shafiq Joty, Alberto Barrón-Cedeño, Giovanni Da San Martino, Simone Filice, Lluís Màrquez, Alessandro Moschitti and Preslav Nakov

Key Concept Identification for Medical Information Retrieval
Jiaping Zheng and Hong Yu

Image-Mediated Learning for Zero-Shot Cross-Lingual Document Retrieval
Ruka Funaki and Hideki Nakayama

15:10–15:40 ***Coffee break***

15:40–17:20 **Session 3A: Sentiment Analysis and Opinion Mining (Long Papers)**

15:40–16:05 *Detecting Risks in the Banking System by Sentiment Analysis*
Clemens Nopp and Allan Hanbury

16:05–16:30 *Sentiment Flow - A General Model of Web Review Argumentation*
Henning Wachsmuth, Johannes Kiesel and Benno Stein

16:30–16:55 *Neural Networks for Open Domain Targeted Sentiment*
Meishan Zhang, Yue Zhang and Duy Tin Vo

16:55–17:20 *Extracting Condition-Opinion Relations Toward Fine-grained Opinion Mining*
Yuki Nakayama and Atsushi Fujii

Saturday, September 19, 2015 (continued)

15:40–17:20 Session 3B: Semantics (Long +TACL Papers)

15:40–16:05 *A large annotated corpus for learning natural language inference*
Samuel R. Bowman, Gabor Angeli, Christopher Potts and Christopher D. Manning

16:05–16:30 *Question-Answer Driven Semantic Role Labeling: Using Natural Language to Annotate Natural Language*
Luheng He, Mike Lewis and Luke Zettlemoyer

16:30–16:55 *[TACL] It's All Fun and Games until Someone Annotates: Video Games with a Purpose for Linguistic Annotation.*
David Jurgens and Roberto Navigli

16:55–17:20 *[TACL] Semantic Proto-Roles*
Drew Reisinger, Rachel Rudinger, Francis Ferraro, Kyle Rawlins and Benjamin Van Durme

15:40–17:20 Session 3C: Information Retrieval and Question Answering (Long Papers)

15:40–16:05 *Name List Only? Target Entity Disambiguation in Short Texts*
Yixin Cao, Juanzi Li, Xiaofei Guo, Shuanhu Bai, Heng Ji and Jie Tang

16:05–16:30 *Biography-Dependent Collaborative Entity Archiving for Slot Filling*
Yu Hong, Xiaobin Wang, Yadong Chen, Jian Wang, Tongtao Zhang and Heng Ji

16:30–16:55 *Stochastic Top-k ListNet*
Tianyi Luo, Dong Wang, Rong Liu and Yiqiao Pan

16:55–17:20 *Exploring Markov Logic Networks for Question Answering*
Tushar Khot, Niranjan Balasubramanian, Eric Gribkoff, Ashish Sabharwal, Peter Clark and Oren Etzioni

Saturday, September 19, 2015 (continued)

15:40–17:20 Session 3D (P1-9): Information Extraction (Long + TACL Paper Posters)

Language and Domain Independent Entity Linking with Quantified Collective Validation
Han Wang, Jin Guang Zheng, Xiaogang Ma, Peter Fox and Heng Ji

Modeling Relation Paths for Representation Learning of Knowledge Bases
Yankai Lin, Zhiyuan Liu, Huanbo Luan, Maosong Sun, Siwei Rao and Song Liu

Corpus-level Fine-grained Entity Typing Using Contextual Information
Yadollah Yaghoobzadeh and Hinrich Schütze

Knowledge Base Unification via Sense Embeddings and Disambiguation
Claudio Delli Bovi, Luis Espinosa Anke and Roberto Navigli

Open-Domain Name Error Detection using a Multi-Task RNN
Hao Cheng, Hao Fang and Mari Ostendorf

Extracting Relations between Non-Standard Entities using Distant Supervision and Imitation Learning
Isabelle Augenstein, Andreas Vlachos and Diana Maynard

Sieve-Based Spatial Relation Extraction with Expanding Parse Trees
Jennifer D'Souza and Vincent Ng

[TACL] Cross-Document Co-Reference Resolution using Sample-Based Clustering with Knowledge Enrichment
Sourav Dutta and Gerhard Weikum

[TACL] Combining Minimally-supervised Methods for Arabic Named Entity Recognition
Maha Althobaiti, Udo Kruschwitz and Massimo Poesio

Saturday, September 19, 2015 (continued)

15:40–17:20 Session 3E (P1-13): Text Mining and NLP Applications (Short Paper Posters)

Mr. Bennet, his coachman, and the Archbishop walk into a bar but only one of them gets recognized: On The Difficulty of Detecting Characters in Literary Texts
Hardik Vala, David Jurgens, Andrew Piper and Derek Ruths

Convolutional Sentence Kernel from Word Embeddings for Short Text Categorization
Jonghoon Kim, Francois Rousseau and Michalis Vazirgiannis

Predicting the Structure of Cooking Recipes
Jermsak Jermsurawong and Nizar Habash

TSDPMM: Incorporating Prior Topic Knowledge into Dirichlet Process Mixture Models for Text Clustering
Linmei Hu, Juanzi Li, Xiaoli Li, Chao Shao and Xuzhong Wang

Sentence Modeling with Gated Recursive Neural Network
Xinchi Chen, Xipeng Qiu, Chenxi Zhu, Shiyu Wu and Xuanjing Huang

Learning Timeline Difference for Text Categorization
Fumiyo Fukumoto and Yoshimi Suzuki

Summarizing Topical Contents from PubMed Documents Using a Thematic Analysis
Sun Kim, Lana Yeganova and W John Wilbur

Recognizing Biographical Sections in Wikipedia
Alessio Palmero Aprosio and Sara Tonelli

Learn to Solve Algebra Word Problems Using Quadratic Programming
Lipu Zhou, Shuaixiang Dai and Liwei Chen

An Unsupervised Method for Discovering Lexical Variations in Roman Urdu Informal Text
Abdul Rafae, Abdul Qayyum, Muhammad Moeenuddin, Asim Karim, Hassan Sajjad and Faisal Kamiran

Component-Enhanced Chinese Character Embeddings
Yanran Li, Wenjie Li, Fei Sun and Sujian Li

Multi-label Text Categorization with Joint Learning Predictions-as-Features Method
Li Li, Houfeng Wang, Xu Sun, Baobao Chang, Shi Zhao and Lei Sha

A Framework for Comparing Groups of Documents
Arun Maiya

Saturday, September 19, 2015 (continued)

Sunday, September 20, 2015

07:30–18:00 Registration

09:00–10:00 Session P2: Plenary Session

Invited Talk: Measuring How Elected Officials and Constituents Communicate
Justin Grimmer

10:00–10:30 ***Coffee break***

10:30–12:10 Session 4A: Information Extraction (Long Papers)

10:30–10:55 *C3EL: A Joint Model for Cross-Document Co-Reference Resolution and Entity Linking*
Sourav Dutta and Gerhard Weikum

10:55–11:20 *Joint Mention Extraction and Classification with Mention Hypergraphs*
Wei Lu and Dan Roth

11:20–11:45 *FINET: Context-Aware Fine-Grained Named Entity Typing*
Luciano Del Corro, Abdalghani Abujabal, Rainer Gemulla and Gerhard Weikum

11:45–12:10 *Joint Entity Recognition and Disambiguation*
Gang Luo, Xiaojiang Huang, Chin-Yew Lin and Zaiqing Nie

10:30–12:10 Session 4B: Statistical Models and Machine Learning Methods (Long Papers)

10:30–10:55 *How Much Information Does a Human Translator Add to the Original?*
Barret Zoph, Marjan Ghazvininejad and Kevin Knight

10:55–11:20 *Hierarchical Recurrent Neural Network for Document Modeling*
Rui Lin, Shujie Liu, Muyun Yang, Mu Li, Ming Zhou and Sheng Li

11:20–11:45 *Auto-Sizing Neural Networks: With Applications to n-gram Language Models*
Kenton Murray and David Chiang

11:45–12:10 *Dual Decomposition Inference for Graphical Models over Strings*
Nanyun Peng, Ryan Cotterell and Jason Eisner

Sunday, September 20, 2015 (continued)

10:30–12:10 Session 4C: Discourse (Long +TACL Papers)

10:30–10:55 *Discourse parsing for multi-party chat dialogues*
Stergos Afantenos, Eric Kow, Nicholas Asher and Jérémy Perret

10:55–11:20 *Joint prediction in MST-style discourse parsing for argumentation mining*
Andreas Peldszus and Manfred Stede

11:20–11:45 *[TACL] One Vector is Not Enough: Entity-Augmented Distributed Semantics for Discourse Relations*
Yangfeng Ji and Jacob Eisenstein

11:45–12:10 *[TACL] Latent Structures for Coreference Resolution*
Sebastian Martschat and Michael Strube

10:30–12:10 Session 4D (P1-9): Semantics (Long Paper Posters)

Feature-Rich Two-Stage Logistic Regression for Monolingual Alignment
Md Arafat Sultan, Steven Bethard and Tamara Sumner

Semantic Role Labeling with Neural Network Factors
Nicholas FitzGerald, Oscar Täckström, Kuzman Ganchev and Dipanjan Das

RELLY: Inferring Hypernym Relationships Between Relational Phrases
Adam Grycner, Gerhard Weikum, Jay Pujara, James Foulds and Lise Getoor

Mise en Place: Unsupervised Interpretation of Instructional Recipes
Chloé Kiddon, Ganesa Thandavam Ponnuraj, Luke Zettlemoyer and Yejin Choi

Semantic Framework for Comparison Structures in Natural Language
Omid Bakhshandeh and James Allen

Sarcastic or Not: Word Embeddings to Predict the Literal or Sarcastic Meaning of Words
Debanjan Ghosh, Weiwei Guo and Smaranda Muresan

Incorporating Trustiness and Collective Synonym/Contrastive Evidence into Taxonomy Construction
Tuan Luu Anh, Jung-jae Kim and See Kiong Ng

Learning to Automatically Solve Logic Grid Puzzles
Arindam Mitra and Chitta Baral

Sunday, September 20, 2015 (continued)

10:30–12:10 Session 4E (P1-13): Machine Translation and Multilinguality (Short Paper Posters)

Improving fast_align by Reordering
Chenchen Ding, Masao Utiyama and Eiichiro Sumita

Touch-Based Pre-Post-Editing of Machine Translation Output
Benjamin Marie and Aurélien Max

A Discriminative Training Procedure for Continuous Translation Models
Quoc-Khanh DO, Alexandre Allauzen and François Yvon

System Combination for Machine Translation through Paraphrasing
Wei-Yun Ma and Kathleen McKeown

Hierarchical Incremental Adaptation for Statistical Machine Translation
Joern Wuebker, Spence Green and John DeNero

ReVal: A Simple and Effective Machine Translation Evaluation Metric Based on Recurrent Neural Networks
Rohit Gupta, Constantin Orasan and Josef van Genabith

Investigating Continuous Space Language Models for Machine Translation Quality Estimation
Kashif Shah, Raymond W. M. Ng, Fethi Bougares and Lucia Specia

Supervised Phrase Table Triangulation with Neural Word Embeddings for Low-Resource Languages
Tomer Levinboim and David Chiang

Translation Invariant Word Embeddings
Kejun Huang, Matt Gardner, Evangelos Papalexakis, Christos Faloutsos, Nikos Sidiropoulos, Tom Mitchell, Partha P. Talukdar and Xiao Fu

Hierarchical Phrase-based Stream Decoding
Andrew Finch, Xiaolin Wang, Masao Utiyama and Eiichiro Sumita

Rule Selection with Soft Syntactic Features for String-to-Tree Statistical Machine Translation
Fabienne Braune, Nina Seemann and Alexander Fraser

Motivating Personality-aware Machine Translation
Shachar Mirkin, Scott Nowson, Caroline Brun and Julien Perez

Trans-gram, Fast Cross-lingual Word-embeddings
Jocelyn Coulmance, Jean-Marc Marty, Guillaume Wenzek and Amine Benhalloum

Sunday, September 20, 2015 (continued)

10:30–12:10 Session 4E (P14-16): Computational Psycholinguistics (Short Paper Posters)

The Overall Markedness of Discourse Relations
Lifeng Jin and Marie-Catherine de Marneffe

Experiments in Open Domain Deception Detection
Verónica Pérez-Rosas and Rada Mihalcea

A model of rapid phonotactic generalization
Tal Linzen and Timothy O'Donnell

12:10–12:50 *Lunch*

12:50–13:30 Session P3: SIGDAT business meeting

13:30–15:10 Session 5A: Text Mining and NLP Applications (Long + TACL Papers)

13:30–13:55 *[TACL] Unsupervised Identification of Translationese*
Ella Rabinovich and Shuly Wintner

13:55–14:20 *Automatically Solving Number Word Problems by Semantic Parsing and Reasoning*
Shuming Shi, Yuehui Wang, Chin-Yew Lin, Xiaojiang Liu and Yong Rui

14:20–14:45 *[TACL] Which Step Do I Take First? Troubleshooting with Bayesian Models*
Annie Louis and Mirella Lapata

14:45–15:10 *[TACL] Problems in Current Text Simplification Research: New Data Can Help*
Wei Xu, Chris Callison-Burch and Courtney Napoles

Sunday, September 20, 2015 (continued)

13:30–15:10 Session 5B: Semantics (Long +TACL Papers)

13:30–13:55 *Parsing English into Abstract Meaning Representation Using Syntax-Based Machine Translation*
Michael Pust, Ulf Hermjakob, Kevin Knight, Daniel Marcu and Jonathan May

13:55–14:20 *The Forest Convolutional Network: Compositional Distributional Semantics with a Neural Chart and without Binarization*
Phong Le and Willem Zuidema

14:20–14:45 *Alignment-Based Compositional Semantics for Instruction Following*
Jacob Andreas and Dan Klein

14:45–15:10 *[TACL] Context-aware Frame-Semantic Role Labeling*
Michael Roth and Mirella Lapata

13:30–15:10 Session 5C: Phonology and Word Segmentation (Long Papers)

13:30–13:55 *Do we need bigram alignment models? On the effect of alignment quality on transduction accuracy in G2P*
Steffen Eger

13:55–14:20 *Keyboard Logs as Natural Annotations for Word Segmentation*
Fumihiko Takahasi and Shinsuke Mori

14:20–14:45 *Long Short-Term Memory Neural Networks for Chinese Word Segmentation*
Xinchi Chen, Xipeng Qiu, Chenxi Zhu, Pengfei Liu and Xuanjing Huang

14:45–15:10 *Semi-supervised Chinese Word Segmentation based on Bilingual Information*
Wei Chen and Bo Xu

13:30–15:10 Session 5D (P1-8): Machine Translation and Multilinguality (Long Paper Posters)

Hierarchical Back-off Modeling of Hiero Grammar based on Non-parametric Bayesian Model
Hidetaka Kamigaito, Taro Watanabe, Hiroya Takamura, Manabu Okumura and Eiichiro Sumita

Consistency-Aware Search for Word Alignment
Shiqi Shen, Yang Liu, Maosong Sun and Huanbo Luan

Graph-Based Collective Lexical Selection for Statistical Machine Translation
Jinsong Su, Deyi Xiong, Shujian Huang, Xianpei Han and Junfeng Yao

Bilingual Correspondence Recursive Autoencoder for Statistical Machine Translation
Jinsong Su, Deyi Xiong, Biao Zhang, Yang Liu, Junfeng Yao and Min Zhang

How to Avoid Unwanted Pregnancies: Domain Adaptation using Neural Network Models
Shafiq Joty, Hassan Sajjad, Nadir Durrani, Kamla Al-Mannai, Ahmed Abdelali and Stephan Vogel

Detecting Content-Heavy Sentences: A Cross-Language Case Study
Junyi Jessy Li and Ani Nenkova

Search-Aware Tuning for Hierarchical Phrase-based Decoding
Feifei Zhai, Liang Huang and Kai Zhao

Part-of-speech Taggers for Low-resource Languages using CCA Features
Young-Bum Kim, Benjamin Snyder and Ruhi Sarikaya

Sunday, September 20, 2015 (continued)

13:30–15:10 Session 5E (P1-12): Tagging, Syntax and Parsing (Short Paper Posters)

An Improved Tag Dictionary for Faster Part-of-Speech Tagging
Robert Moore

Improving Arabic Diacritization through Syntactic Analysis
Anas Shahrour, Salam Khalifa and Nizar Habash

Combining Discrete and Continuous Features for Deterministic Transition-based Dependency Parsing
Meishan Zhang and Yue Zhang

Efficient Inner-to-outer Greedy Algorithm for Higher-order Labeled Dependency Parsing
Xuezhe Ma and Eduard Hovy

Online Updating of Word Representations for Part-of-Speech Tagging
Wenpeng Yin, Tobias Schnabel and Hinrich Schütze

Empty Category Detection using Path Features and Distributed Case Frames
Shunsuke Takeno, Masaaki Nagata and Kazuhide Yamamoto

Foreebank: Syntactic Analysis of Customer Support Forums
Rasoul Kaljahi, Jennifer Foster, Johann Roturier, Corentin Ribeyre, Teresa Lynn and Joseph Le Roux

Semi-supervised Dependency Parsing using Bilexical Contextual Features from Auto-Parsed Data
Eliyahu Kiperwasser and Yoav Goldberg

Improved Transition-Based Parsing and Tagging with Neural Networks
Chris Alberti, David Weiss, Greg Coppola and Slav Petrov

Syntactic Parse Fusion
Do Kook Choe, David McClosky and Eugene Charniak

Not All Contexts Are Created Equal: Better Word Representations with Variable Attention
Wang Ling, Yulia Tsvetkov, Silvio Amir, Ramon Fermandez, Chris Dyer, Alan W Black, Isabel Trancoso and Chu-Cheng Lin

An Improved Non-monotonic Transition System for Dependency Parsing
Matthew Honnibal and Mark Johnson

Sunday, September 20, 2015 (continued)

15:10–15:40 ***Coffee break***

15:40–17:20 **Session 6A: Machine Translation (Long Papers)**

15:40–16:05 *Improving Statistical Machine Translation with a Multilingual Paraphrase Database*
Ramtin Mehdizadeh Seraj, Maryam Siahbani and Anoop Sarkar

16:05–16:30 *Learning Semantic Representations for Nonterminals in Hierarchical Phrase-Based Translation*
Xing Wang, Deyi Xiong and Min Zhang

16:30–16:55 *A Comparison between Count and Neural Network Models Based on Joint Translation and Reordering Sequences*
Andreas Guta, Tamer Alkhouli, Jan-Thorsten Peter, Joern Wuebker and Hermann Ney

16:55–17:20 *Effective Approaches to Attention-based Neural Machine Translation*
Thang Luong, Hieu Pham and Christopher D. Manning

15:40–17:20 **Session 6B: Sentiment Analysis and Opinion Mining / Tagging, Chunking and Parsing (Long Papers)**

15:40–16:05 *Document Modeling with Gated Recurrent Neural Network for Sentiment Classification*
Duyu Tang, Bing Qin and Ting Liu

16:05–16:30 *Fine-grained Opinion Mining with Recurrent Neural Networks and Word Embeddings*
Pengfei Liu, Shafiq Joty and Helen Meng

16:30–16:55 *Joint A* CCG Parsing and Semantic Role Labelling*
Mike Lewis, Luheng He and Luke Zettlemoyer

16:55–17:20 *Improving Semantic Parsing with Enriched Synchronous Context-Free Grammar*
Junhui Li, Muhua Zhu, Wei Lu and Guodong Zhou

Sunday, September 20, 2015 (continued)

15:40–17:20 Session 6C: Language and Vision / Information Extraction (Long Papers)

15:40–16:05 *Solving Geometry Problems: Combining Text and Diagram Interpretation*
Minjoon Seo, Hannaneh Hajishirzi, Ali Farhadi, Oren Etzioni and Clint Malcolm

16:05–16:30 *Do You See What I Mean? Visual Resolution of Linguistic Ambiguities*
Yevgeni Berzak, Andrei Barbu, Daniel Harari, Boris Katz and Shimon Ullman

16:30–16:55 *Efficient and Expressive Knowledge Base Completion Using Subgraph Feature Extraction*
Matt Gardner and Tom Mitchell

16:55–17:20 *Representing Text for Joint Embedding of Text and Knowledge Bases*
Kristina Toutanova, Danqi Chen, Patrick Pantel, Hoifung Poon, Pallavi Choudhury and Michael Gamon

15:40–17:20 Session 6D (P1-11): Statistical Models and Machine Learning Methods (Long Paper Posters)

A Utility Model of Authors in the Scientific Community
Yanchuan Sim, Bryan Routledge and Noah A. Smith

Finding Function in Form: Compositional Character Models for Open Vocabulary Word Representation
Wang Ling, Chris Dyer, Alan W Black, Isabel Trancoso, Ramon Fermandez, Silvio Amir, Luis Marujo and Tiago Luis

Syntax-Aware Multi-Sense Word Embeddings for Deep Compositional Models of Meaning
Jianpeng Cheng and Dimitri Kartsaklis

Conversation Trees: A Grammar Model for Topic Structure in Forums
Annie Louis and Shay B. Cohen

Fast, Flexible Models for Discovering Topic Correlation across Weakly-Related Collections
Jingwei Zhang, Aaron Gerow, Jaan Altosaar, James Evans and Richard Jean So

Molding CNNs for text: non-linear, non-consecutive convolutions
Tao Lei, Regina Barzilay and Tommi Jaakkola

Multi-Perspective Sentence Similarity Modeling with Convolutional Neural Networks
Hua He, Kevin Gimpel and Jimmy Lin

Posterior calibration and exploratory analysis for natural language processing models
Khanh Nguyen and Brendan O'Connor

A Generative Word Embedding Model and its Low Rank Positive Semidefinite Solution
Shaohua Li, Jun Zhu and Chunyan Miao

Reading Documents for Bayesian Online Change Point Detection
Taehoon Kim and Jaesik Choi

15:40–17:20 Session 6E (P1-13): Semantics (Short Paper Posters)

Recognizing Textual Entailment Using Probabilistic Inference
Lei Sha, Sujian Li, Baobao Chang, Zhifang Sui and Tingsong Jiang

Chinese Semantic Role Labeling with Bidirectional Recurrent Neural Networks
Zhen Wang, Tingsong Jiang, Baobao Chang and Zhifang Sui

Unsupervised Negation Focus Identification with Word-Topic Graph Model
Bowei Zou, Guodong Zhou and Qiaoming Zhu

Reverse-engineering Language: A Study on the Semantic Compositionality of German Compounds
Corina Dima

Event Detection and Factuality Assessment with Non-Expert Supervision
Kenton Lee, Yoav Artzi, Yejin Choi and Luke Zettlemoyer

Large-Scale Acquisition of Entailment Pattern Pairs by Exploiting Transitivity
Julien Kloetzer, Kentaro Torisawa, Chikara Hashimoto and Jong-Hoon Oh

Context-Dependent Knowledge Graph Embedding
Yuanfei Luo, Quan Wang, Bin Wang and Li Guo

Learning to Identify the Best Contexts for Knowledge-based WSD
Evgenia Wasserman Pritsker, William Cohen and Einat Minkov

Sunday, September 20, 2015 (continued)

Measuring Prerequisite Relations Among Concepts
Chen Liang, Zhaohui Wu, Wenyi Huang and C. Lee Giles

Adapting Phrase-based Machine Translation to Normalise Medical Terms in Social Media Messages
Nut Limsopatham and Nigel Collier

Script Induction as Language Modeling
Rachel Rudinger, Pushpendre Rastogi, Francis Ferraro and Benjamin Van Durme

Online Learning of Interpretable Word Embeddings
Hongyin Luo, Zhiyuan Liu, Huanbo Luan and Maosong Sun

A Strong Lexical Matching Method for the Machine Comprehension Test
Ellery Smith, Nicola Greco, Matko Bosnjak and Andreas Vlachos

19:00–23:00 Conference Dinner

Monday, September 21, 2015

07:30–18:00 Registration

09:00–10:00 Session P4: Plenary Session

09:00–09:05 *Best Paper Awards*
Chris Callison-Burch and Jian Su

09:05–09:30 *Broad-coverage CCG Semantic Parsing with AMR*
Yoav Artzi, Kenton Lee and Luke Zettlemoyer

09:30–09:55 *Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems*
Tsung-Hsien Wen, Milica Gasic, Nikola Mrkšić, Pei-Hao Su, David Vandyke and Steve Young

09:55–10:05 *A large annotated corpus for learning natural language inference*
Samuel R. Bowman, Gabor Angeli, Christopher Potts and Christopher D. Manning

10:05–10:30 ***Coffee break***

Monday, September 21, 2015 (continued)

10:30–12:10 Session 7A: Semantics (Long +TACL Papers)

10:30–10:55 *Do Multi-Sense Embeddings Improve Natural Language Understanding?*
Jiwei Li and Dan Jurafsky

10:55–11:20 *Learning Semantic Composition to Detect Non-compositionality of Multiword Expressions*
Majid Yazdani, Meghdad Farahmand and James Henderson

11:20–11:45 *Solving General Arithmetic Word Problems*
Subhro Roy and Dan Roth

11:45–12:10 *[TACL] From Paraphrase Database to Compositional Paraphrase Model and Back*
John Wieting, Mohit Bansal, Kevin Gimpel, Karen Livescu and Dan Roth

10:30–12:10 Session 7B: Information Extraction (Long Papers)

10:30–10:55 *Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks*
Daojian Zeng, Kang Liu, Yubo Chen and Jun Zhao

10:55–11:20 *CORE: Context-Aware Open Relation Extraction with Factorization Machines*
Fabio Petroni, Luciano Del Corro and Rainer Gemulla

11:20–11:45 *Improved Relation Extraction with Feature-Rich Compositional Embedding Models*
Matthew R. Gormley, Mo Yu and Mark Dredze

11:45–12:10 *Classifying Relations via Long Short Term Memory Networks along Shortest Dependency Paths*
Yan Xu, Lili Mou, Ge Li, Yunchuan Chen, Hao Peng and Zhi Jin

10:30–12:10 Session 7C: Computational Psycholinguistics / Machine Translation (Long Papers)

10:30–10:55 *A Computational Cognitive Model of Novel Word Generalization*
Aida Nematzadeh, Erin Grant and Suzanne Stevenson

10:55–11:20 *Personality Profiling of Fictional Characters using Sense-Level Links between Lexical Resources*
Lucie Flekova and Iryna Gurevych

11:20–11:45 *Leave-one-out Word Alignment without Garbage Collector Effects*
Xiaolin Wang, Masao Utiyama, Andrew Finch, Taro Watanabe and Eiichiro Sumita

11:45–12:10 *Generalized Agreement for Bidirectional Word Alignment*
Chunyang Liu, Yang Liu, Maosong Sun, Huanbo Luan and Heng Yu

10:30–12:10 Session 7D (P1-6): Word Segmentation, Tagging and Parsing (Long +TACL Paper Posters)

A Transition-based Model for Joint Segmentation, POS-tagging and Normalization
Tao Qian, Yue Zhang, Meishan Zhang, Yafeng Ren and Donghong Ji

Multilingual discriminative lexicalized phrase structure parsing
Benoit Crabbé

Hierarchical Low-Rank Tensors for Multilingual Transfer Parsing
Yuan Zhang and Regina Barzilay

Diversity in Spectral Learning for Natural Language Parsing
Shashi Narayan and Shay B. Cohen

Transition-based Dependency Parsing Using Two Heterogeneous Gated Recursive Neural Networks
Xinchi Chen, Yaqian Zhou, Chenxi Zhu, Xipeng Qiu and Xuanjing Huang

[TACL] A Graph-based Lattice Dependency Parser for Joint Morphological Segmentation and Syntactic Analysis
Wolfgang Seeker and Özlem Çetinoğlu

10:30–12:10 Session 7E (P1-3): Spoken Language Processing (Short Paper Posters)

Turn-taking phenomena in incremental dialogue systems
Hatim Khouzaimi, Romain Laroche and Fabrice Lefevre

Hierarchical Latent Words Language Models for Robust Modeling to Out-Of Domain Tasks
Ryo Masumura, Taichi Asami, Takanobu Oba, Hirokazu Masataki, Sumitaka Sakauchi and Akinori Ito

A Coarse-Grained Model for Optimal Coupling of ASR and SMT Systems for Speech Translation
Gaurav Kumar, Graeme Blackwood, Jan Trmal, Daniel Povey and Sanjeev Khudanpur

10:30–12:10 Session 7E (P4-18): Summarization (Short Paper Posters)

Abstractive Multi-document Summarization with Semantic Information Extraction
Wei Li

Concept-based Summarization using Integer Linear Programming: From Concept Pruning to Multiple Optimal Solutions
Florian Boudin, Hugo Mougard and Benoit Favre

GhostWriter: Using an LSTM for Automatic Rap Lyric Generation
Peter Potash, Alexey Romanov and Anna Rumshisky

Better Summarization Evaluation with Word Embeddings for ROUGE
Jun-Ping Ng and Viktoria Abrecht

Krimping texts for better summarization
Marina Litvak, Mark Last and Natalia Vanetik

From the Virtual to the RealWorld: Referring to Objects in Real-World Spatial Scenes
Dimitra Gkatzia, Verena Rieser, Phil Bartie and William Mackaness

An Unsupervised Bayesian Modelling Approach for Storyline Detection on News Articles
Deyu Zhou, Haiyang Xu and Yulan He

Topical Coherence for Graph-based Extractive Summarization
Daraksha Parveen, Hans-Martin Ramsl and Michael Strube

Monday, September 21, 2015 (continued)

Summarizing Student Responses to Reflection Prompts
Wencan Luo and Diane Litman

Extractive Summarization by Maximizing Semantic Volume
Dani Yogatama, Fei Liu and Noah A. Smith

LCSTS: A Large Scale Chinese Short Text Summarization Dataset
Baotian Hu, Qingcai Chen and Fangze Zhu

Discourse Planning with an N-gram Model of Relations
Or Biran and Kathleen McKeown

Experiments with Generative Models for Dependency Tree Linearization
Richard Futrell and Edward Gibson

Summarization Based on Embedding Distributions
Hayato Kobayashi, Masaki Noguchi and Taichi Yatsuka

Reversibility reconsidered: finite-state factors for efficient probabilistic sampling in parsing and generation
Marc Dymetman, Sriram Venkatapathy and Chunyang Xiao

12:10–13:30 ***Lunch***

13:30–15:15 **Session 8A: Fun and Quirky Topics (Short Papers)**

13:30–13:45 *A quantitative analysis of gender differences in movies using psycholinguistic normatives*
Anil Ramakrishna, Nikolaos Malandrakis, Elizabeth Staruk and Shrikanth Narayanan

13:45–14:00 *EMNLP versus ACL: Analyzing NLP research over time*
Sujatha Das Gollapalli and Xiaoli Li

14:00–14:15 *Answering Elementary Science Questions by Constructing Coherent Scenes using Background Knowledge*
Yang Li and Peter Clark

14:15–14:30 *WikiQA: A Challenge Dataset for Open-Domain Question Answering*
Yi Yang, Wen-tau Yih and Christopher Meek

Monday, September 21, 2015 (continued)

14:30–14:45 *Personalized Machine Translation: Predicting Translational Preferences*
Shachar Mirkin and Jean-Luc Meunier

14:45–15:00 *Talking to the crowd: What do people react to in online discussions?*
Aaron Jaech, Victoria Zayats, Hao Fang, Mari Ostendorf and Hannaneh Hajishirzi

15:00–15:15 *What Your Username Says About You*
Aaron Jaech and Mari Ostendorf

13:30–15:15 Session 8B: Semantics (Short Papers)

13:30–13:45 *Knowledge Base Inference using Bridging Entities*
Bhushan Kotnis, Pradeep Bansal and Partha P. Talukdar

13:45–14:00 *Specializing Word Embeddings for Similarity or Relatedness*
Douwe Kiela, Felix Hill and Stephen Clark

14:00–14:15 *Evaluation of Word Vector Representations by Subspace Alignment*
Yulia Tsvetkov, Manaal Faruqui, Wang Ling, Guillaume Lample and Chris Dyer

14:15–14:30 *Higher-order logical inference with compositional semantics*
Koji Mineshima, Pascual Martínez-Gómez, Yusuke Miyao and Daisuke Bekki

14:30–14:45 *Any-language frame-semantic parsing*
Anders Johannsen, Héctor Martínez Alonso and Anders Søgaard

15:00–15:15 *What's in an Embedding? Analyzing Word Embeddings through Multilingual Evaluation*
Arne Köhn

13:30–15:15 Session 8C: Statistical Models and Machine Learning Methods, Machine Translation (Short Papers)

13:30–13:45 *Joint Event Trigger Identification and Event Coreference Resolution with Structured Perceptron*
Jun Araki and Teruko Mitamura

13:45–14:00 *A Joint Dependency Model of Morphological and Syntactic Structure for Statistical Machine Translation*
Rico Sennrich and Barry Haddow

14:00–14:15 *Variable-Length Word Encodings for Neural Translation Models*
Rohan Chitnis and John DeNero

14:15–14:30 *A Binarized Neural Network Joint Model for Machine Translation*
Jingyi Zhang, Masao Utiyama, Eiichiro Sumita, Graham Neubig and Satoshi Nakamura

14:30–14:45 *Bayesian Optimization of Text Representations*
Dani Yogatama, Lingpeng Kong and Noah A. Smith

14:45–15:00 *A Comparative Study on Regularization Strategies for Embedding-based Neural Networks*
Hao Peng, Lili Mou, Ge Li, Yunchuan Chen, Yangyang Lu and Zhi Jin

15:00–15:15 *Efficient Hyper-parameter Optimization for NLP Applications*
Lidan Wang, Minwei Feng, Bowen Zhou, Bing Xiang and Sridhar Mahadevan

13:30–15:15 Session 8D (P1-6): NLP for Web and Social Media, including Computational Social Science (Long Paper Posters)

Improved Arabic Dialect Classification with Social Media Data
Fei Huang

Exploiting Debate Portals for Semi-Supervised Argumentation Mining in User-Generated Web Discourse
Ivan Habernal and Iryna Gurevych

Confounds and Consequences in Geotagged Twitter Data
Umashanthi Pavalanathan and Jacob Eisenstein

Modeling Reportable Events as Turning Points in Narrative
Jessica Ouyang and Kathleen McKeown

Towards the Extraction of Customer-to-Customer Suggestions from Reviews
Sapna Negi and Paul Buitelaar

Using Content-level Structures for Summarizing Microblog Repost Trees
Jing Li, Wei Gao, Zhongyu Wei, Baolin Peng and Kam-Fai Wong

13:30–15:15 Session 8D (P7-9): Discourse (Long Paper Posters)

Intra-sentential Zero Anaphora Resolution using Subject Sharing Recognition
Ryu Iida, Kentaro Torisawa, Chikara Hashimoto, Jong-Hoon Oh and Julien Kloetzer

Estimation of Discourse Segmentation Labels from Crowd Data
Ziheng Huang, Jialu Zhong and Rebecca J. Passonneau

Comparing Word Representations for Implicit Discourse Relation Classification
Chloé Braud and Pascal Denis

13:30–15:15 Session 8E (P1-9): Discourse (Short Paper Posters)

Better Document-level Sentiment Analysis from RST Discourse Parsing
Parminder Bhatia, Yangfeng Ji and Jacob Eisenstein

Closing the Gap: Domain Adaptation from Explicit to Implicit Discourse Relations
Yangfeng Ji, Gongbo Zhang and Jacob Eisenstein

Wikification of Concept Mentions within Spoken Dialogues Using Domain Constraints from Wikipedia
Seokhwan Kim, Rafael E. Banchs and Haizhou Li

Shallow Convolutional Neural Network for Implicit Discourse Relation Recognition
Biao Zhang, Jinsong Su, Deyi Xiong, Yaojie Lu, Hong Duan and Junfeng Yao

On the Role of Discourse Markers for Discriminating Claims and Premises in Argumentative Discourse
Judith Eckle-Kohler, Roland Kluge and Iryna Gurevych

Fatal or not? Finding errors that lead to dialogue breakdowns in chat-oriented dialogue systems
Ryuichiro Higashinaka, Masahiro Mizukami, Kotaro Funakoshi, Masahiro Araki, Hiroshi Tsukahara and Yuka Kobayashi

Learning Word Meanings and Grammar for Describing Everyday Activities in Smart Environments
Muhammad Attamimi, Yuji Ando, Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi and Hideki Asoh

Discourse Element Identification in Student Essays based on Global and Local Cohesion
Wei Song, Ruiji Fu, Lizhen Liu and Ting Liu

Adapting Coreference Resolution for Narrative Processing
Quynh Ngoc Thi Do, Steven Bethard and Marie-Francine Moens

13:30–15:15 Session 8E (P10-15): Phonology, Morphology and Word Segmentation (Short Paper Posters)

Joint Lemmatization and Morphological Tagging with Lemming
Thomas Müller, Ryan Cotterell, Alexander Fraser and Hinrich Schütze

Transducer Disambiguation with Sparse Topological Features
Gonzalo Iglesias, Adrià de Gispert and Bill Byrne

Arabic Diacritization with Recurrent Neural Networks
Yonatan Belinkov and James Glass

Automatic Diacritics Restoration for Hungarian
Attila Novák and Borbála Siklósi

Morphological Analysis for Unsegmented Languages using Recurrent Neural Network Language Model
Hajime Morita, Daisuke Kawahara and Sadao Kurohashi

Can Symbol Grounding Improve Low-Level NLP? Word Segmentation as a Case Study
Hirotaka Kameko, Shinsuke Mori and Yoshimasa Tsuruoka

15:15–15:40 ***Coffee break***

Monday, September 21, 2015 (continued)

15:40–17:20 Session 9A: Statistical Models and Machine Learning Methods (Long + TACL Papers)

15:40–16:05 *When Are Tree Structures Necessary for Deep Learning of Representations?*
Jiwei Li, Thang Luong, Dan Jurafsky and Eduard Hovy

16:05–16:30 *Discriminative Neural Sentence Modeling by Tree-Based Convolution*
Lili Mou, Hao Peng, Ge Li, Yan Xu, Lu Zhang and Zhi Jin

16:30–16:55 *Multi-Timescale Long Short-Term Memory Neural Network for Modelling Sentences and Documents*
Pengfei Liu, Xipeng Qiu, Xinchi Chen, Shiyu Wu and Xuanjing Huang

16:55–17:20 *[TACL] Learning Structural Kernels for Natural Language Processing*
Daniel Beck, Trevor Cohn, Christian Hardmeier and Lucia Specia

15:40–17:20 Session 9B: Text Mining and NLP Applications (Long Papers)

15:40–16:05 *Verbal and Nonverbal Clues for Real-life Deception Detection*
Verónica Pérez-Rosas, Mohamed Abouelenien, Rada Mihalcea, Yao Xiao, CJ Linton and Mihai Burzo

16:05–16:30 *Social Media Text Classification under Negative Covariate Shift*
Geli Fei and Bing Liu

16:30–16:55 *Co-Training for Topic Classification of Scholarly Data*
Cornelia Caragea, Florin Bulgarov and Rada Mihalcea

16:55–17:20 *Humor Recognition and Humor Anchor Extraction*
Diyi Yang, Alon Lavie, Chris Dyer and Eduard Hovy

Monday, September 21, 2015 (continued)

15:40–17:20 Session 9C: Spoken Language Processing and Language Modeling (Long Papers)

15:40–16:05 *Topic Identification and Discovery on Text and Speech*
Chandler May, Francis Ferraro, Alan McCree, Jonathan Wintrode, Daniel Garcia-Romero and Benjamin Van Durme

16:05–16:30 *A Dynamic Programming Algorithm for Computing N-gram Posteriors from Lattices*
Dogan Can and Shrikanth Narayanan

16:30–16:55 *Bilingual Structured Language Models for Statistical Machine Translation*
Ekaterina Garmash and Christof Monz

16:55–17:20 *Compact, Efficient and Unlimited Capacity: Language Modeling with Compressed Suffix Trees*
Ehsan Shareghi, Matthias Petri, Gholamreza Haffari and Trevor Cohn

15:40–17:20 Session 9D (P1-8): Semantics (Long Paper Posters)

ERSOM: A Structural Ontology Matching Approach Using Automatically Learned Entity Representation
Chuncheng Xiang, Tingsong Jiang, Baobao Chang and Zhifang Sui

A Single Word is not Enough: Ranking Multiword Expressions Using Distributional Semantics
Martin Riedl and Chris Biemann

Syntactic Dependencies and Distributed Word Representations for Analogy Detection and Mining
Likun Qiu, Yue Zhang and Yanan Lu

Navigating the Semantic Horizon using Relative Neighborhood Graphs
Amaru Cuba Gyllensten and Magnus Sahlgren

Multi- and Cross-Modal Semantics Beyond Vision: Grounding in Auditory Perception
Douwe Kiela and Stephen Clark

Automatic recognition of habituals: a three-way classification of clausal aspect
Annemarie Friedrich and Manfred Pinkal

Distributed Representations for Unsupervised Semantic Role Labeling
Kristian Woodsend and Mirella Lapata

A Tableau Prover for Natural Logic and Language
Lasha Abzianidze

15:40–17:20 Session 9E (P1-9): Sentiment Analysis and Opinion Mining (Short Paper Posters)

JEAM: A Novel Model for Cross-Domain Sentiment Classification Based on Emotion Analysis
Kun-Hu Luo, Zhi-Hong Deng, Hongliang Yu and Liang-Chen Wei

PhraseRNN: Phrase Recursive Neural Network for Aspect-based Sentiment Analysis
Thien Hai Nguyen and Kiyoaki Shirai

ASTD: Arabic Sentiment Tweets Dataset
Mahmoud Nabil, Mohamed Aly and Amir Atiya

Adjective Intensity and Sentiment Analysis
Raksha Sharma, Mohit Gupta, Astha Agarwal and Pushpak Bhattacharyya

The Rating Game: Sentiment Rating Reproducibility from Text
Lasse Borgholt, Peter Simonsen and Dirk Hovy

A Multi-lingual Annotated Dataset for Aspect-Oriented Opinion Mining
Salud M. Jiménez-Zafra, Giacomo Berardi, Andrea Esuli, Diego Marcheggiani, María Teresa Martín-Valdivia and Alejandro Moreo Fernández

Deep Convolutional Neural Network Textual Features and Multiple Kernel Learning for Utterance-level Multimodal Sentiment Analysis
Soujanya Poria, Erik Cambria and Alexander Gelbukh

SLSA: A Sentiment Lexicon for Standard Arabic
Ramy Eskander and Owen Rambow

Reinforcing the Topic of Embeddings with Theta Pure Dependence for Text Classification
Ning Xing, Yuexian Hou, Peng Zhang, Wenjie Li and Dawei Song

15:40–17:20 **Session 9E (P10-17): NLP for Web and Social Media, including Computational Social Science (Short Paper Posters)**

That's So Annoying!!!: A Lexical and Frame-Semantic Embedding Based Data Augmentation Approach to Automatic Categorization of Annoying Behaviors using #petpeeve Tweets
William Yang Wang and Diyi Yang

Detection of Steganographic Techniques on Twitter
Alex Wilson, Phil Blunsom and Andrew Ker

#SupportTheCause: Identifying Motivations to Participate in Online Health Campaigns
Dong Nguyen, Tijs van den Broek, Claudia Hauff, Djoerd Hiemstra and Michel Ehrenhard

An Analysis of Domestic Abuse Discourse on Reddit
Nicolas Schrading, Cecilia Ovesdotter Alm, Ray Ptucha and Christopher Homan

Twitter-scale New Event Detection via K-term Hashing
Dominik Wurzer, Victor Lavrenko and Miles Osborne

Classifying Tweet Level Judgements of Rumours in Social Media
Michal Lukasik, Trevor Cohn and Kalina Bontcheva

Identification and Verification of Simple Claims about Statistical Properties
Andreas Vlachos and Sebastian Riedel

17:30–17:50 **Session P5: Closing Remarks**

18:30–20:00 **Farewell Drink**

List of Papers

Knowledge Base Unification via Sense Embeddings and Disambiguation

Claudio Delli Bovi
Department of Computer Science
Sapienza University of Rome
dellibovi@di.uniroma1.it

Luis Espinosa-Anke
Department of Information and Communication Technologies
Universitat Pompeu Fabra
luis.espinosa@upf.edu

Roberto Navigli
Department of Computer Science
Sapienza University of Rome
navigli@di.uniroma1.it

Abstract

We present KB-UNIFY, a novel approach for integrating the output of different Open Information Extraction systems into a single unified and fully disambiguated knowledge repository. KB-UNIFY consists of three main steps: (1) disambiguation of relation argument pairs via a sense-based vector representation and a large unified sense inventory; (2) ranking of semantic relations according to their degree of specificity; (3) cross-resource relation alignment and merging based on the semantic similarity of domains and ranges. We tested KB-UNIFY on a set of four heterogeneous knowledge bases, obtaining high-quality results. We discuss and provide evaluations at each stage, and release output and evaluation data for the use and scrutiny of the community[1].

1 Introduction

The breakthrough of the Open Information Extraction (OIE) paradigm opened up a research area where Web-scale unconstrained Information Extraction systems are developed to acquire and formalize large quantities of knowledge. However, while successful, to date most state-of-the-art OIE systems have been developed with their own type inventories, and no portable ontological structure. In fact, OIE systems can be very different in nature. Early approaches (Etzioni et al., 2008; Wu and Weld, 2010; Fader et al., 2011) focused on extracting a large number of relations from massive unstructured corpora, mostly relying on dependencies at the level of surface text. Systems like NELL (Carlson et al., 2010) combine a hand-crafted taxonomy of entities and relations with self-supervised large-scale extraction from the Web, but they require additional processing for linking and integration (Dutta et al., 2014).

More recent work has focused, instead, on deeper language understanding, especially at the level of syntax and semantics (Nakashole et al., 2012; Moro and Navigli, 2013). By leveraging semantic analysis, knowledge gathered from unstructured text can be adequately integrated and used to enrich existing knowledge bases, such as YAGO (Mahdisoltani et al., 2015), FREEBASE (Bollacker et al., 2008) and DBPEDIA (Lehmann et al., 2014). A large amount of reliable structured knowledge is crucial for OIE approaches based on distant supervision (Mintz et al., 2009; Riedel et al., 2010), even when multi-instance multi-learning algorithms (Surdeanu et al., 2012) or matrix factorization techniques (Riedel et al., 2013; Fan et al., 2014) come into play to deal with noisy extractions. For this reason a recent trend of research has focused on Knowledge Base (KB) completion (Nickel et al., 2012; Bordes et al., 2013), exploiting the fact that distantly supervised OIE and structured knowledge can complement each other. However, the majority of integration approaches nowadays are not designed to deal with many different resources at the same time.

We propose an approach where the key idea is to bring together knowledge drawn from an arbitrary number of OIE systems, regardless of whether these systems provide links to some general-purpose inventory, come with their own ad-hoc structure, or have no structure at all. Knowledge from each source, in the form of ⟨subject, predicate, object⟩ triples, is disambiguated and linked to a single large sense inventory. This enables us to discover alignments at a semantic level between relations from different KBs, and to generate a unified, fully disambiguated KB of entities and semantic relations. KB-UNIFY achieves state-of-the-art disambiguation and provides a general, resource-independent representation of semantic relations, suitable for any kind of KB.

[1] http://lcl.uniroma1.it/kb-unify

Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 726–736,
Lisbon, Portugal, 17-21 September 2015. ©2015 Association for Computational Linguistics.

The remainder of this paper is structured as follows: Section 2 reviews relevant related work; Sections 3, 4, 5 and 6 describe in detail each stage of the approach; Sections 7 and 8 describe the experiments carried out and the results obtained; and finally Section 9 summarizes our findings and discusses potential directions for future work.

2 Related Work

The integration of knowledge drawn from different sources has received much attention over the last decade. Among the most notable examples are resources like BabelNet (Navigli and Ponzetto, 2012), UBY (Gurevych et al., 2012) and YAGO (Mahdisoltani et al., 2015). While great effort has been put into aligning knowledge at the *concept level*, most approaches do not tackle the problem of integrating heterogeneous knowledge at the *relation level*, nor do they exploit effectively the huge amount of information harvested with OIE systems, even when this information is unambiguously linked to a structured resource, as in (Nakashole et al., 2012), or (Moro and Navigli, 2013). In fact, as the number of resources increases, KB alignment is already becoming an emergent research field: Dutta et al. (2014) describe a method for linking arguments in NELL triples to DBPEDIA by combining First Order Logic and Markov Networks; Grycner and Weikum (2014) semantify PATTY's pattern synsets and connect them to WordNet verbs; Lin et al. (2012) propose a method to propagate FREEBASE types across REVERB and deal with the problem of unlinkable entities. All these approaches achieve very competitive results in their respective settings, but unlike the approach being proposed here, they limit the task to 1-to-1 alignments. A few contributions have tried to broaden the scope and include different resources at the same time, but with rather different goals from ours. For example, Riedel et al. (2013) propose a universal schema that integrates structured data with OIE data by learning latent feature vectors for entities and relations; the KNOWLEDGE VAULT (Dong et al., 2014) uses a graph-based probabilistic framework where prior knowledge from existing resources (e.g. FREEBASE) improves Web extractions by predicting their reliability. However, in both cases the main objective is distantly supervised extraction from unstructured text, rather than KB unification. A recent trend of research focuses on learning embedding models for structured knowledge and their application to tasks like relation extraction and KB completion (Socher et al., 2013; Weston et al., 2013; Bordes et al., 2013). These approaches, however, leverage embeddings at surface level, which are suboptimal for our task, as will be discussed in Section 3. Since we require a common semantic framework for KB unification, we use vector representations based on word senses, which are mapped to a very large sense inventory. This shared sense inventory, then, constitutes the common ground in which disambiguation, alignment and final unification occurs.

3 Knowledge Base Unification: Overview

KB-UNIFY takes as input a set of KBs $\mathbf{K} = \{KB_1, ..., KB_n\}$ and outputs a single, unified and fully disambiguated KB, denoted as $\mathbf{KB}^*$. For our purposes we can define a KB KB_i as a triple $\langle E_i, R_i, T_i \rangle$, where E_i is a set of entities, R_i is a set of semantic relations, and T_i is a set of triples (facts) $\langle e_d, r, e_g \rangle$ with subject and object e_d, $e_g \in E_i$ and predicate $r \in R_i$. Depending on the nature of each KB_i, entities in E_i might be disambiguated and linked to an external inventory (e.g. the entity *Washington* linked to the Wikipedia page GEORGE WASHINGTON), or unlinked and only available as ambiguous mentions (e.g. the bare word *washington* might refer to the president, the city or the state). We can thus partition $\mathbf{K}$ into a subset of linked resources $\mathbf{K}_D$, and one of unlinked resources $\mathbf{K}_U$. In order to align very different and heterogeneous KBs at the semantic level, KB-UNIFY exploits:

- A unified sense inventory S, which acts as a superset for the inventories of individual KBs. We choose BabelNet (Navigli and Ponzetto, 2012) for this purpose: by merging complementary knowledge from different resources (e.g. Wikipedia, WordNet, Wikidata and Wiktionary, among others), BabelNet provides a wide coverage of entities and concepts whilst at the same time enabling convenient inter-resource mappings for KB_i in $\mathbf{K}_D$. For instance, each Wikipedia page (or Wikidata item) has a corresponding synset in BabelNet, which enables a one-to-one mapping between BabelNet's synsets and entries in, e.g., DBPEDIA or FREEBASE;

- A vector space model V_S that enables a semantic representation for every item in S. Current distributional models, like word em-

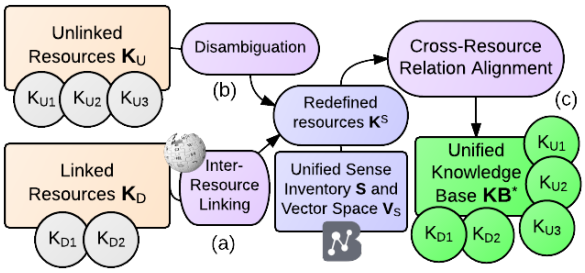

Figure 1: Unification algorithm workflow

beddings (Mikolov et al., 2013), are not suitable to our setting: they are constrained to surface word forms, and hence they inherently retain ambiguity of polysemous words and entity mentions. We thus leverage SENSEMBED (Iacobacci et al., 2015), a novel semantically-enhanced approach to embeddings. SENSEMBED is trained on a large annotated corpus and produces continuous representations for individual word senses (*sense embeddings*), according to an underlying sense inventory.

Figure 1 illustrates the workflow of our KB unification approach. Entities coming from any $KB_i \in \mathbf{K}_D$ can be directly (and unambiguously) mapped to the corresponding entries in S via BabelNet inter-resource linking (Figure 1(a)): in the above example, the entity *Washington* linked to the Wikipedia page GEORGE WASHINGTON is included in the BabelNet synset $\texttt{Washington}^4_{bn}$. In contrast, unlinked (and potentially ambiguous) entities need an explicit disambiguation step (Figure 1(b)) connecting them to appropriate entries, i.e. synsets, in S: this is the case, in the above example, for the ambiguous mention *washington* that has to be linked to either the president, the city or the state. Therefore, our approach comprises two successive stages:

- A **disambiguation** stage (Section 5) where all $KB_i \in \mathbf{K}$ are linked to S, either by inter-resource mapping (Figure 1(a)) or disambiguation (Figure 1(b)), and all E_i are merged into a unified set of entities E^*. As a result of this process we obtain a set $\mathbf{K}^S$ comprising all the KBs in $\mathbf{K}$ redefined using the common sense inventory S;
- An **alignment** stage (Section 6, Figure 1(c)) where, for each pair of KBs $KB_i^S, KB_j^S \in \mathbf{K}^S$, we compare any relation pair $\langle r_i, r_j \rangle$, $r_i \in R_i^S$ and $r_j \in R_j^S$, in order to identify cross-resource alignments and merge relations sharing equivalent semantics into relation clusters (*relation synsets*). This process yields a unified set of relation synsets R^*. The overall result is $\mathbf{KB}^* = \langle E^*, R^*, T^* \rangle$, where T^* is the set of all disambiguated triples redefined over E^* and R^*.

4 Background

The disambiguation stage of our approach is based on the interplay between two core components: a vector space model V_S, as introduced in Section 3, which provides an unambiguous semantic representation for each item in S; and a Word Sense Disambiguation/Entity Linking system, working on the same sense inventory S, which discovers and disambiguates concepts and entity mentions within a given input text. In this section we briefly describe our choice for these two components: SENSEMBED (Iacobacci et al., 2015) and BABELFY (Moro et al., 2014).

SENSEMBED is a knowledge-based approach for obtaining latent continuous representations of individual word senses. Unlike other sense-based embeddings approaches, like (Huang et al., 2012), which address the inherent polysemy of word-level representations relying solely on text corpora, SENSEMBED exploits the structured knowledge of a large sense inventory along with the distributional information gathered from text corpora. In order to do this, SENSEMBED requires a sense-annotated corpus; for each target word sense, then, a representation is computed by maximizing the log likelihood of the word sense with respect to its context within the annotated text, similarly to the word-based embeddings model. Following Iacobacci et al. (2015), we trained SENSEMBED using the English Wikipedia and, as sense inventory, BabelNet.

BABELFY[2] is a joint state-of-the-art approach to multilingual Entity Linking and Word Sense Disambiguation. Given the BabelNet lexicalized semantic network as underlying structure, BABELFY first models each concept in the network through its corresponding *semantic signature* by leveraging a graph random walk algorithm. Then, given an input text, the generated semantic signatures are used to construct a subgraph

[2] http://babelfy.org

of the semantic network representing the meaning of the content words in that text. BABELFY then searches this subgraph for the intended sense of each content word, by means of a densest-subgraph heuristic that identifies high-coherence interpretations. Given its unified approach that covers concepts and named entities alike, and its flexibility in disambiguating both bag-of-words and proper text, BABELFY constitutes the most convenient choice for linking relation triples to a high-coverage sense inventory like BabelNet.

5 Disambiguation

In the disambiguation phase (Figure 1(b)), all $KB_i \in \mathbf{K}_U$ are linked to the unified sense inventory S and added to the set of redefined KBs $\mathbf{K}^S$. As explained in Section 3, while each KB in $\mathbf{K}_D$ can be unambiguously redefined via BabelNet inter-resource links and added to $\mathbf{K}^S$, KBs in $\mathbf{K}_U$ require an explicit disambiguation step. Given $KB_i \in \mathbf{K}_U$, our disambiguation module (Figure 2) takes as input its set of unlinked triples T_i and outputs a set $T_i^S \subseteq T_i$ of disambiguated triples with subject-object pairs linked to S. The triples in T_i^S, together with their corresponding entity sets and relation sets, constitute the redefined KB_i^S which is then added to $\mathbf{K}^S$. However, applying a straightforward approach that disambiguates all triples in isolation might lead to very imprecise results, due to the lack of available context for each individual triple. We thus devised a disambiguation strategy that comprises three successive steps:

1. We identify a set of high-confidence seeds from T_i (Section 5.1), i.e. triples $\langle e_d, r, e_g \rangle$ where subject e_d and object e_g are highly semantically related, and disambiguate them using the senses that maximize their similarity in our vector space V_S;

2. We use the seeds to generate a ranking of the relations in R_i according to their degree of specificity (Section 5.2). We represent each $r \in R_i$ in our vector space V_S and assign higher specificity to relations whose arguments are closer in V_S;

3. We finally disambiguate the remaining non-seed triples in T_i (Section 5.3) starting from the most specific relations, and jointly using all participating argument pairs as context.

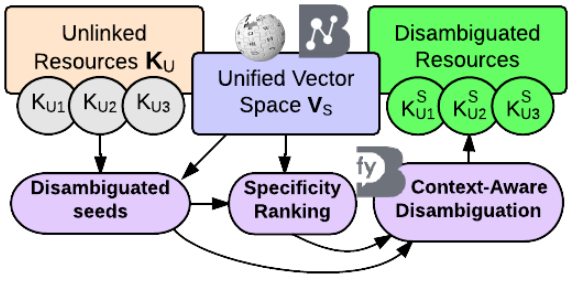

Figure 2: Disambiguation algorithm workflow

5.1 Identifying Seed Argument Pairs

The first stage of our disambiguation approach aims at extracting reliable seeds from T_i, i.e. triples $\langle e_d, r, e_g \rangle$ where subject e_d and object e_g can be confidently disambiguated without additional context. In order to do this we leverage the sense embeddings associated with each candidate disambiguation for e_d and e_g. We consider all the available senses for both e_d and e_g in S, namely $\mathbf{s}_d = \{s_d^1, ..., s_d^m\}$ and $\mathbf{s}_g = \{s_g^1, ..., s_g^{m'}\}$, and the corresponding sets of sense embeddings $\mathbf{v}_d = \{v_d^1, ..., v_d^m\}$ and $\mathbf{v}_g = \{v_g^1, ..., v_g^{m'}\}$. We then select, among all possible pairs of senses, the pair $\langle s_d^*, s_g^* \rangle$ that maximizes the cosine similarity between the corresponding embeddings $\langle v_d^*, v_g^* \rangle$:

$$\langle v_d^*, v_g^* \rangle = \text{argmax}_{v_d \in \mathbf{v}_d,\, v_g \in \mathbf{v}_g} \frac{v_d \cdot v_g}{\|v_d\| \, \|v_g\|} \quad (1)$$

For each disambiguated triple $\langle s_d^*, r, s_g^* \rangle$, the cosine similarity value associated with $\langle v_d^*, v_g^* \rangle$ represents the disambiguation confidence ζ_{dis}. We rank all such triples according to their confidence, and select those above a given threshold δ_{dis}. The underlying assumption is that, for high-confidence subject-object pairs, the embeddings associated with the correct senses s_d^* and s_g^* will be closest in V_S compared to any other candidate pair. Intuitively, the more the relation r between e_d and e_g is semantically well defined, the more this assumption is justified. As an example, consider the triple $\langle$*Armstrong*, *worked for*, *NASA*$\rangle$: among all the possible senses for *Armstrong* (the astronaut, the jazz musician the cyclist, etc.) and *NASA* (the space agency, the racing organization, a Swedish band, etc.) we expect the vectors corresponding to the astronaut and the space agency to be closest in the vector space model V_S.

5.2 Relation Specificity Ranking

The assumption that, given an ambiguous subject-object pair, correct argument senses are

the closest pair in the vector space (Section 5.1) is easily verifiable for general relations (e.g. *is a*, *is part of*). However, as a semantic relation becomes specific, its arguments are less guaranteed to be semantically related (e.g. *is a professor in the university of*) and a disambiguation approach based exclusively on similarity is prone to errors. On the other hand, specific relations tend to narrow down the scope of possible entity types occurring as subject and object. In the above example, *is a professor in the university of* requires entity pairs with professors as subjects and cities as objects. Our disambiguation strategy should thus vary according to the specificity of the relations taken into account. In order to consider this observation in our disambiguation pipeline, we first need to estimate the degree of specificity for each relation in the relation set R_i of the target KB to be disambiguated. Given R_i and a set of seeds from the previous stage (Section 5.1), we apply a specificity ranking policy and sort relations in R_i from the most general to the most specific. We compute the generality $Gen(r)$ of a given relation r by looking at the spatial dispersion of the sense embeddings associated with its seed subjects and objects. Let $\mathbf{v}_D$ ($\mathbf{v}_G$) be the set of sense embeddings associated with the domain (range) seed arguments of r. For both $\mathbf{v}_D$ and $\mathbf{v}_G$, we compute the corresponding centroid vectors μ_D and μ_G as:

$$\mu_k = \frac{1}{|\mathbf{v}_k|} \sum_{v \in \mathbf{v}_k} \frac{v}{\|v\|}, \quad k \in \{D, G\} \qquad (2)$$

Then, the variances σ_D^2 and σ_G^2 are given by:

$$\sigma_k^2 = \frac{1}{|\mathbf{v}_k|} \sum_{v \in \mathbf{v}_k} (1 - cos\,(v, \mu_k))^2 \qquad (3)$$

with $k \in \{D, G\}$ as before. We finally compute $Gen(r)$ as the average of σ_D^2 and σ_G^2. The result of this procedure is a *relation specificity ranking* that associates each relation r with its generality $Gen(r)$. Intuitively, we expect more general relations to show higher variance (hence higher $Gen(r)$), as their subjects and objects are likely to be rather disperse throughout the vector space; instead, arguments of very specific relations are more likely to be clustered together in compact regions, yielding lower values of $Gen(r)$.

5.3 Disambiguation with Relation Context

In the third step, both the specificity ranking and the seeds are exploited to disambiguate the remaining triples in T_i. To do this we leverage BABELFY (Moro et al., 2014) (introduced in Section 4). As we observed in Section 5.2, specific relations impose constraints on their subject-object types and tend to show compact domains and ranges in the vector space. Therefore, given a triple $\langle e_d, r, e_g \rangle$, knowing that r is specific enables us to put together all the triples in T_i where r occurs, and use them to provide meaningful context for disambiguation. If r is general, instead, its subject-object types are less constrained and additional triples do not guarantee to provide semantically related context.

At this stage, our algorithm takes as input the set of triples T_i, along with the associated disambiguation seeds (Section 5.1), the specificity ranking (Section 5.2) and a specificity threshold δ_{spec}. T_i is first partitioned into two subsets: T_i^{spec}, comprising all the triples for which $Gen(r) < \delta_{spec}$, and $T_i^{gen} = T_i \setminus T_i^{spec}$. We then employ two different disambiguation strategies:

- For each distinct relation r occurring in T_i^{spec}, we first retrieve the subset $T_{i,r}^{spec} \subset T_i^{spec}$ of triples where r occurs, and then disambiguate $T_{i,r}^{spec}$ as a whole with BABELFY. For each triple in $T_{i,r}^{spec}$, context is provided by all the remaining triples along with the disambiguated seeds extracted for r.

- We disambiguate the remaining triples in T_i^{gen} one by one in isolation with BABELFY, providing for each triple only the predicate string r as additional context.

6 Cross-Resource Relation Alignment

After disambiguation (Section 5) each KB in $\mathbf{K}$ is linked to the unified sense inventory S and added to $\mathbf{K}^S$. However, each $KB_i^S \in \mathbf{K}^S$ still provides its own relation set $R_i^S \subseteq R_i$. Instead, in the unified $\mathbf{KB}^*$, relations with equivalent semantics should be considered as part of a single relation synset even when they come from different KBs. Therefore, at this stage, we apply an alignment algorithm to identify pairs of relations from different KBs having equivalent semantics. We exploit the fact that each relation r is now defined over entity pairs linked to S, and we generate a semantic representation of r in the vector space V_S based on the centroid vectors of its domain and range. Due to representing the semantics of relations on this common ground, we can compare them by computing their domain and range similarity in V_S. We

first consider each $KB_i^S \in \mathbf{K}^S$ and, for each relation r_i in R_i^S, we compute the corresponding centroid vectors $\mu_d^{r_i}$ and $\mu_g^{r_i}$ using formula (2). Then, for each pair of KBs $\langle KB_i^S, KB_j^S \rangle \in \mathbf{K}^S \times \mathbf{K}^S$, we compare all relation pairs $\langle r_i, r_j \rangle \in R_i^S \times R_j^S$ by computing the cosine similarity between domain centroids s_D and between range centroids s_G:

$$s_k = \frac{\mu_k^{r_i} \cdot \mu_k^{r_j}}{\|\mu_k^{r_i}\| \, \|\mu_k^{r_j}\|} \quad (4)$$

where μ_k^r denotes the centroid associated with relation r and $k \in \{D, G\}$. The average of s_D and s_G gives us an *alignment confidence* ζ_{align} for the pair $\langle r_i, r_j \rangle$. If confidence is above a given threshold δ_{align} then r_i and r_j are merged into the same relation synset. Relations for which no alignment is found are turned into singleton relation synsets. As a result of this alignment procedure we obtain the unified set of relations R^*.

7 Experimental Setup

The setting for our experimental evaluation was the following:

- We used BabelNet 3.0[3] as our unified sense inventory for the unification procedure as well as the underlying inventory for both BABELFY and SENSEMBED. Currently, BabelNet contains around 14M synsets and represents the largest single multilingual repository of entities and concepts;
- We selected PATTY (Nakashole et al., 2012) and WISENET (Moro and Navigli, 2013) as linked resources. We used PATTY with FREEBASE types and pattern synsets derived from Wikipedia, and WISENET 2.0 with Wikipedia relational phrases;
- We selected NELL (Carlson et al., 2010) and REVERB (Fader et al., 2011) as unlinked resources. We used KB beliefs updated to November 2014 for the former, and the set of relation instances from ClueWeb09 for the latter.

Comparative statistics in Table 1 show that the input KBs are rather different in nature: NELL is based on 298 predefined relations and contains beliefs for about 2 million entities. The distribution of entities over relations is however very skewed, with 80.33% of the triples being instances of the `generalizations` relationship. In contrast, REVERB contains a highly sparse relation set (1,299,844 distinct relations) and more than 3 million distinct entities. PATTY features the largest (and, together with WISENET, sparsest) set of triples, with 1,631,531 distinct relations and less than 10 triples per relation on average.

[3] `http://babelnet.org`

	$\mathbf{K}_U$		$\mathbf{K}_D$	
	NELL	REVERB	PATTY	WISENET
# relations	298	1 299 844	1 631 531	245 935
# triples	2 245 050	14 728 268	15 802 946	2 271 807
# entities	1 996 021	3 327 425	1 087 907	1 636 307

Table 1: Statistics on the input KBs

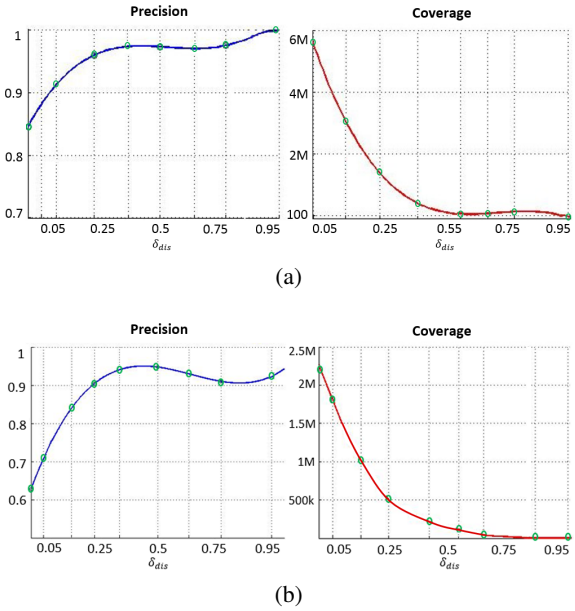

Figure 3: Precision (left) and coverage (right) of disambiguated seeds at different values of δ_{dis} for (a) the whole set of triples in PATTY and (b) the subset of ambiguous triples

8 Experiments

8.1 Disambiguation

We tested our disambiguation approach experimentally in terms of both disambiguated seed quality (Section 8.1.1) and overall disambiguation performance (Section 8.1.2). We created a development set by extracting a subset of 6 million triples from the largest linked KB in our experimental setup, i.e. PATTY. Triples in PATTY are automatically linked to YAGO, which is in turn linked to WordNet and DBPEDIA. Since both resources are also linked by BabelNet, we mapped the original triples to the BabelNet sense inventory and used them to tune our disambiguation module. We also provide two baseline approaches: (1) di-

	SENSEMBED			Baseline		
ζ_{dis}	0.5-0.7	0.7-0.9	0.9-1.0	0.5-0.7	0.7-0.9	0.9-1.0
PATTY	.980	.980	1.000	.793	.780	1.000
WISENET	.958	.960	.973	.726	.786	.791
NELL	.955	.995	1.000	.800	.770	.885
REVERB	.930	.940	.950	.775	.725	.920

Table 2: Disambiguation precision for all KBs

	$\delta_{spec} = 0.8$		$\delta_{spec} = 0.5$		$\delta_{spec} = 0.3$	
	all	only seeds	all	only seeds	all	only seeds
PATTY	62.15	26.60	52.49	24.06	40.75	21.41
WISENET	60.00	37.46	54.44	22.26	53.58	16.62
NELL	76.97	62.98	50.95	20.71	44.70	4.36
REVERB	41.20	38.57	25.14	23.70	13.37	12.75

Table 3: Coverage results (%) for all KBs

rect disambiguation on individual triples with BABELFY alone (without the seeds) and (2) direct disambiguation of the seeds only (without BABELFY).

8.1.1 Results: Disambiguated Seeds

We tuned our disambiguation algorithm by studying the quality of the disambiguated seeds (Section 5.1) extracted from the surface text triples of PATTY. Figure 3 shows precision and coverage for increasing values of the confidence threshold δ_{dis}. We computed precision by checking each disambiguated seed against the corresponding linked triple in the development set, and coverage as the ratio of covered triples. We analyzed results for both the whole set of triples in PATTY (Fig. 3a) and the subset of ambiguous triples (Fig. 3b), i.e. those triples whose subjects and objects have at least two candidate senses each in the BabelNet inventory. In both cases, precision of disambiguated seeds increases rapidly with δ_{dis}, stabilizing above 90% with $\delta_{dis} > 0.25$. Coverage displays the opposite behavior, decreasing exponentially with more confident outcomes, from 6 million triples to less than a thousand (for seeds with confidence $\delta_{dis} > 0.95$). As a result, we chose $\delta_{dis} = 0.25$ as optimal threshold value throughout the rest of the evaluations.

In addition, we manually evaluated the disambiguated seeds extracted from both linked KBs (PATTY and WISENET) and unlinked KBs (NELL and REVERB). For each KB, we extracted up to three random samples of 150 triples according to different levels of confidence ζ_{dis}: the first sample included extraction with $0.5 \leq \zeta_{dis} < 0.7$, the second with $0.7 \leq \zeta_{dis} < 0.9$, and the third with $\zeta_{dis} \geq 0.9$. Each sample was evaluated by two human judges: for each disambiguated triple

	KB-UNIFY		Dutta et al.	Baseline
	all	only seeds	($\alpha = 0.5$)	
Precision	.852	**.957**	.931	.749
Recall	**.875**	.117	.799	.608
F-score	**.864**	.197	.857	.671

Table 4: Disambiguation results over NELL gold standard

$\langle e_d, r, e_g \rangle$, we presented our judges with the surface text arguments e_d, e_g and the relation string r, along with the two BabelNet synsets corresponding to the disambiguated arguments s_d^*, s_g^*, and we asked whether the association of each subject and object with the proposed BabelNet synset was correct. We then estimated precision as the average proportion of correctly disambiguated triples. For each sample we compared disambiguation precision using SENSEMBED, as in Section 5.1, against the first baseline with BABELFY alone. Results, reported in Table 2, show that our approach consistently outperforms the baseline and provides high precision over all samples and KBs.

8.1.2 Results: Disambiguation with Relation Context

We then evaluated the overall disambiguation output after specificity ranking (Section 5.2) and disambiguation with relation context using BABELFY (Section 5.3). We analyzed three configurations of the disambiguation pipeline, namely $\delta_{spec} \in \{0.8, 0.5, 0.3\}$. We ran the algorithm over both linked and unlinked KBs of our experimental setup, and computed the coverage for each KB as the overall ratio of disambiguated triples. Results are reported in Table 3 and compared to the coverage obtained from the disambiguated seeds only: context-aware disambiguation substantially increases coverage over all KBs. Table 3 also shows that a restrictive δ_{spec} results in lower coverage values, due to the increased number of triples disambiguated without context.

Finally, we evaluated the quality of disambiguation on a publicly available dataset (Dutta et al., 2014) comprising manual annotations for NELL. This dataset provides a gold standard of 1200 triples whose subjects and objects are manually assigned a proper DBpedia URI. We again used BabelNet's inter-resource links to express DBpedia annotations with our sense inventory and then sought, for each annotated triple in the dataset, the corresponding triple in our disambiguated version of NELL with $\delta_{dis} = 0.25$ and $\delta_{spec} = 0.8$. We then repeated this process con-

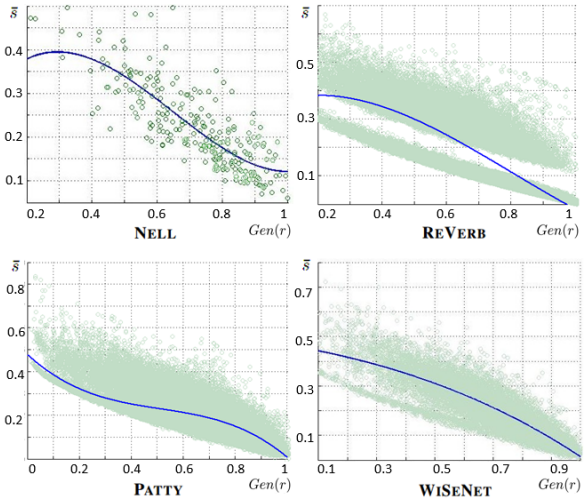

Figure 4: Average argument similarity against $Gen(r)$

	NELL	REVERB	PATTY	WISENET
Precision	.660	.715	.625	.750
Cohen's kappa	-	.430	.620	.600

Table 5: Specificity ranking evaluation

	NELL
High $Gen(r)$	`agent created` `at location`
Low $Gen(r)$	`person in economic sector` `restaurant in city`
	REVERB
High $Gen(r)$	`is for` `is in`
Low $Gen(r)$	`enter Taurus in` `carry oxygen to`
	PATTY
High $Gen(r)$	`located in` `later served to`
Low $Gen(r)$	`starting pitcher who played` `league coach for`
	WISENET
High $Gen(r)$	`include` `is a type of`
Low $Gen(r)$	`lobe-finned fish lived during` `took part in the Eurovision contest`

Table 6: Examples of general and specific relations for all KBs

sidering only the disambiguated seeds instead of the whole disambiguation pipeline. In line with (Dutta et al., 2014), we computed precision, recall and F-score for each setting. Results are reported in Table 4 and compared against those of Dutta et al. (2014) and against our first baseline with BABELFY alone. KB-UNIFY achieves the best result, showing that a baseline based on state-of-the-art disambiguation is negatively affected by the lack of context for each individual triple. In contrast, an approach that relies only on the disambiguated seeds affords very high precision, but suffers from dramatically lower coverage.

8.2 Specificity Ranking

We evaluated the specificity ranking (Section 5.2) generated by KB-UNIFY for all KBs of our experimental setup. First of all, we empirically validated our scoring function $Gen(r)$ over each resource: for each relation we computed the average cosine similarity among all its domain arguments $\bar{s}_D$ and among all its range arguments $\bar{s}_G$. We then plotted the average $\bar{s}$ of $\bar{s}_D$ and $\bar{s}_G$ against $Gen(r)$ for each relation r (Figure 4). As observed in Section 5.2, the average similarity among domain and range arguments decreases for increasing values of $Gen(r)$, indicating that more general relations allow less semantically constrained subject-object types. We then used human judgement to assess the quality of our specificity rankings. First, each ranking was split into four quartiles, and two human evaluators were presented with a sample from the top quartile (i.e. a relation falling into the most general category) and a sample from the bottom quartile (i.e. a relation falling into the most specific category). We shuffled each relation pair, showed it to our human judges, and then asked which of the two relations they considered to be the more specific. Ranking precision was computed by considering those pairs where human choice agreed with the ranking. Finally, we computed inter-annotator agreement on each specificity ranking (except for NELL, due to the small sample size) with Cohen's kappa coefficient (Cohen, 1968). Results for each ranking are reported in Table 5, while some examples of general and specific relations for each KB are shown in Table 6. Disagreement between human choice and ranking is higher in NELL (where the set of relations is quite small compared to other KBs) and in PATTY (due to a sparser set of relations, biased towards very specific patterns). Inter-annotator agreement is instead lower for REVERB, where unconstrained Web harvesting often results in ambiguous relation strings.

8.3 Alignment

Due to the novelty of our approach, and hence the lack of widely accepted gold standards and testbeds, we evaluated our cross-resource relation alignment algorithm (Section 6) by exploiting human judgement once again. Given the results of

	PATTY-WISENET		PATTY-REVERB		NELL-REVERB	
δ_{align}	0.7	0.9	0.7	0.9	0.7	0.9
Prec.	.68	**.80**	.58	**.74**	.61	**.75**
# Align.	128k	1.2k	47k	643	2.6k	88
	PATTY-NELL		**WISENET-NELL**		**WISENET-REVERB**	
δ_{align}	0.7	0.9	0.7	0.9	0.7	0.9
Prec.	.66	**1.00**	.70	**.84**	.59	**.87**
# Align.	2.6k	57	381	34	9.9k	169

Table 7: Cross-resource alignment evaluation

PATTY-WISENET		ζ_{align}
portrayed	's character	0.84
debuted in	first appeared in	0.86
PATTY-REVERB		ζ_{align}
language in	is spoken in	0.81
mostly known for	plays the role of	0.70
NELL-REVERB		ζ_{align}
bookwriter	is a novel by	0.88
personleadscity	is the mayor of	0.60
NELL-PATTY		ζ_{align}
worksfor	was hired by	0.72
riveremptiesintoriver	tributary of	0.89
NELL-WISENET		ζ_{align}
animaleatfood	feeds on	0.72
teamhomestadium	play their home games at	0.88
REVERB-WISENET		ζ_{align}
has a selection of	offers	0.82
had grown up in	was born and raised in	0.85

Table 8: Examples of cross resource relation alignments

Section 8.1, we considered the top 10k frequent relations for each KB and ran the algorithm over each possible pair of KBs with two different configurations: $\delta_{align} = 0.7$ and $\delta_{align} = 0.9$. From each pair of KBs $\langle KB_i, KB_j \rangle$ we obtained a list of candidate alignments, i.e. pairs of relations $\langle r_i, r_j \rangle$ where $r_i \in KB_i$ and $r_j \in KB_j$. From each list we then extracted a random sample of 150 candidate alignments. We showed each alignment[4] $\langle r_i, r_j \rangle$ to two human judges, and asked whether r_i and r_j represented the same relation. The problem was presented in terms of paraphrasing: for each pair, we asked whether exchanging r_i and r_j within a sentence would have changed that sentence's meaning. In line with Section 8.2 we computed precision based on the agreement between human choice and automatic alignments. Results are reported in Table 7. Our alignment algorithm shows high precision in all pairings where $\delta_{align} = 0.9$. Alignment reliability decreases for lower δ_{align}, as relation pairs where r_i is a generalization of r_j (or vice versa) tend to have similar centroids in V_S. The same holds for pairs where r_i is the negation of r_j (or vice versa). Even though we could have utilized measures based on relation string similarity (Dutta et al., 2015) to reduce wrong alignments in these cases, by relying on a purely semantic criterion we removed any prior assumption on the format of input KBs. Some examples of alignments are shown in Table 8.

To conclude, we report statistics regarding the unified **KB*** produced from the initial set of resources in our experimental setup (cf. Section 7). We validated our thresholds for high-precision, and selected $\delta_{dis} = 0.25$, δ_{spec} = 0.8 and δ_{align} = 0.8. Our alignment algorithm produced 56,673 confident alignments, out of which 2,207 relation synsets were derived, with an average size of 16.82 individual relations per synset. As a result, we obtained a unified **KB*** comprising 24,221,856 disambiguated triples defined over 1,952,716 distinct entities and 2,675,296 distinct relations.

9 Conclusion and Future Work

We have presented KB-UNIFY, a novel, general approach for disambiguating and seamlessly unifying KBs produced by different OIE systems. KB-UNIFY represents entities and relations using a shared semantic representation, leveraging a unified sense inventory together with a semantically-enhanced vector space model and a disambiguation algorithm. This enables us to disambiguate unlinked resources (like NELL and REVERB) with high precision and coverage, and to discover relation-level cross-resource alignments effectively. One of the key features of our strategy is its generality: by representing each KB on a common ground, we need no prior assumption on the nature and format of the knowledge it encodes. We tested our approach experimentally on a set of four very different KBs, both linked and unlinked, and we evaluated disambiguation and alignment results extensively at every stage, exploiting both human evaluations and public gold standard datasets (when available). This work opens compelling avenues for future work. We plan to further exploit sense-enhanced unified representations of relations in various ways: providing an ontological structure for the unified KB, exploring complementary approaches for capturing semantic relation alignments, and incorporating multilinguality.

Acknowledgments

The authors gratefully acknowledge the support of the ERC Starting Grant MultiJEDI No. 259234.

[4] In the case of relation synsets, such as PATTY and WISENET, we selected up to three random relation strings from each synset.

References

Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: A Collaboratively Created Graph Database For Structuring Human Knowledge. In *Proceedings of SIGMOD*, pages 1247–1250.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. In *Advances in NIPS*, volume 26, pages 2787–2795.

Andrew Carlson, Justin Betteridge, Bryan Kisiel, Burr Settles, Estevam R. Hruschka Jr., and Tom M. Mitchell. 2010. Toward an Architecture for Never-Ending Language Learning. In *Proceedings of AAAI*, pages 1306–1313.

Jacob Cohen. 1968. Weighted Kappa: Nominal Scale Agreement Provision for Scaled Disagreement or Partial Credit. *Psychological Bulletin*, 70(4):213–220.

Xin Dong, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. 2014. Knowledge Vault: A Web-scale Approach to Probabilistic Knowledge Fusion. In *Proceedings of the SIGKDD*, pages 601–610.

Arnab Dutta, Christian Meilicke, and Simone Paolo Ponzetto. 2014. A Probabilistic Approach for Integrating Heterogeneous Knowledge Sources. In *Proceedings of ESWC*, pages 286–301.

Arnab Dutta, Christian Meilicke, and Heiner Stuckenschmidt. 2015. Enriching Structured Knowledge with Open Information. In *Proceedings of WWW*, pages 267–277.

Oren Etzioni, Michele Banko, Stephen Soderland, and Daniel S. Weld. 2008. Open Information Extraction from the Web. *Commun. ACM*, 51(12):68–74.

Anthony Fader, Stephen Soderland, and Oren Etzioni. 2011. Identifying Relations for Open Information Extraction. In *Proceedings of EMNLP*, pages 1535–1545.

Miao Fan, Deli Zhao, Qiang Zhou, Zhiyuan Liu, Thomas Fang Zheng, and Edward Y. Chang. 2014. Distant Supervision for Relation Extraction with Matrix Completion. In *Proceedings of ACL*, pages 839–849.

Adam Grycner and Gerhard Weikum. 2014. HARPY: Hypernyms and Alignment of Relational Paraphrases. In *Proceedings of ACL*, pages 2195–2204.

Iryna Gurevych, Judith Eckle-Kohler, Silvana Hartmann, Michael Matuschek, Christian M. Meyer, and Christian Wirth. 2012. Uby: A large-scale unified lexical-semantic resource based on LMF. In *Proceedings of ACL*, pages 580–590.

Eric H. Huang, Richard Socher, Christopher D. Manning, and Andrew Y. Ng. 2012. Improving Word Representations via Global Context and Multiple Word Prototypes. In *Proceedings of ACL*, pages 873–882.

Ignacio Iacobacci, Mohammad Taher Pilehvar, and Roberto Navigli. 2015. SensEmbed: Learning Sense Embeddings for Word and Relational Similarity. In *Proceedings of ACL*, pages 95–105.

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. 2014. DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web Journal*, pages 1–29.

Thomas Lin, Mausam, and Oren Etzioni. 2012. No Noun Phrase Left Behind: Detecting and Typing Unlinkable Entities. In *Proceedings of EMNLP-CoNLL*, pages 893–903.

Farzaneh Mahdisoltani, Joanna Biega, and Fabian M. Suchanek. 2015. YAGO3: A Knowledge Base from Multilingual Wikipedias. In *CIDR*.

Tomas Mikolov, Kal Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *Proceedings of Workshop at ICLR*.

Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant Supervision for Relation Extraction Without Labeled Data. In *Proceedings of ACL-IJCNLP*, pages 1003–1011.

Andrea Moro and Roberto Navigli. 2013. Integrating Syntactic and Semantic Analysis into the Open Information Extraction Paradigm. In *Proceedings of IJCAI*, pages 2148–2154.

Andrea Moro, Alessandro Raganato, and Roberto Navigli. 2014. Entity Linking meets Word Sense Disambiguation: a Unified Approach. *TACL*, 2:231–244.

Ndapandula Nakashole, Gerhard Weikum, and Fabian M. Suchanek. 2012. PATTY: A Taxonomy of Relational Patterns with Semantic Types. In *Proceedings of EMNLP-CoNLL*, pages 1135–1145.

Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage Multilingual Semantic Network. *Artificial Intelligence*, 193:217–250.

Maximilian Nickel, Volker Tresp, and Hans-Peter Kriegel. 2012. Factorizing YAGO: Scalable Machine Learning for Linked Data. In *Proceedings of WWW*, pages 271–280.

Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. Modeling Relations and Their Mentions Without Labeled Text. In *Proceedings of ECML-PKDD*, pages 148–163.

Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M. Marlin. 2013. Relation Extraction with Matrix Factorization and Universal Schemas. In *Proceedings of NAACL*, pages 74–84.

Richard Socher, Danqi Chen, Christopher D. Manning, and Andrew Y. Ng. 2013. Reasoning with Neural Tensor Networks for Knowledge Base Completion. In *Advances in NIPS*, pages 926–934.

Mihai Surdeanu, Julie Tibshirani, Ramesh Nallapati, and Christopher D. Manning. 2012. Multi-instance Multi-label Learning for Relation Extraction. In *Proceedings of EMNLP-CoNLL*, pages 455–465.

Jason Weston, Antoine Bordes, Oksana Yakhnenko, and Nicolas Usunier. 2013. Connecting Language and Knowledge Bases with Embedding Models for Relation Extraction. In *Proceedings of EMNLP*, pages 1366–1371.

Fei Wu and Daniel S. Weld. 2010. Open Information Extraction using Wikipedia. In *Proceedings of ACL*, pages 118–127.