The 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing

Proceedings of the Conference
Volume 2: Short Papers

ACL 2015
July 26-31
Beijing, China

Platinum Sponsors:

Gold Sponsors:

Silver Sponsors:

Bronze Sponsor:

Best Paper Sponsor:

©2015 The Association for Computational Linguistics

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

ISBN 978-1-941643-73-0

Table of Contents

A Framework for the Construction of Monolingual and Cross-lingual Word Similarity Datasets
José Camacho-Collados, Mohammad Taher Pilehvar and Roberto Navigli 1

On metric embedding for boosting semantic similarity computations
Julien Subercaze, Christophe Gravier and Frédérique Laforest 8

Improving Distributed Representation of Word Sense via WordNet Gloss Composition and Context Clustering
Tao Chen, Ruifeng Xu, Yulan He and Xuan Wang 15

A Multitask Objective to Inject Lexical Contrast into Distributional Semantics
Nghia The Pham, Angeliki Lazaridou and Marco Baroni 21

Semi-Stacking for Semi-supervised Sentiment Classification
Shoushan Li, Lei Huang, Jingjing Wang and Guodong Zhou 27

Deep Markov Neural Network for Sequential Data Classification
Min Yang 32

Semantic Analysis and Helpfulness Prediction of Text for Online Product Reviews
Yinfei Yang, Yaowei Yan, Minghui Qiu and Forrest Bao 38

Document Classification by Inversion of Distributed Language Representations
Matt Taddy 45

Using Tweets to Help Sentence Compression for News Highlights Generation
Zhongyu Wei, Yang Liu, Chen Li and Wei Gao 50

Domain-Specific Paraphrase Extraction
Ellie Pavlick, Juri Ganitkevitch, Tsz Ping Chan, Xuchen Yao, Benjamin Van Durme and Chris Callison-Burch 57

Simplifying Lexical Simplification: Do We Need Simplified Corpora?
Goran Glavaš and Sanja Štajner 63

Zoom: a corpus of natural language descriptions of map locations
Romina Altamirano, Thiago Ferreira, Ivandré Paraboni and Luciana Benotti 69

Generating overspecified referring expressions: the role of discrimination
Ivandré Paraboni, Michelle Galindo and Douglas Iacovelli 76

Using prosodic annotations to improve coreference resolution of spoken text
Ina Roesiger and Arndt Riester 83

Spectral Semi-Supervised Discourse Relation Classification
Robert Fisher and Reid Simmons 89

I do not disagree: leveraging monolingual alignment to detect disagreement in dialogue
Ajda Gokcen and Marie-Catherine de Marneffe 94

Language Models for Image Captioning: The Quirks and What Works
Jacob Devlin, Hao Cheng, Hao Fang, Saurabh Gupta, Li Deng, Xiaodong He, Geoffrey Zweig and Margaret Mitchell 100

A Distributed Representation Based Query Expansion Approach for Image Captioning
Semih Yagcioglu, Erkut Erdem, Aykut Erdem and Ruket Cakici 106

Learning language through pictures
Grzegorz Chrupała, Ákos Kádár and Afra Alishahi 112

Exploiting Image Generality for Lexical Entailment Detection
Douwe Kiela, Laura Rimell, Ivan Vulić and Stephen Clark 119

Lexicon Stratification for Translating Out-of-Vocabulary Words
Yulia Tsvetkov and Chris Dyer 125

Recurrent Neural Network based Rule Sequence Model for Statistical Machine Translation
Heng Yu and Xuan Zhu 132

Discriminative Preordering Meets Kendall's τ Maximization
Sho Hoshino, Yusuke Miyao, Katsuhito Sudoh, Katsuhiko Hayashi and Masaaki Nagata 139

Evaluating Machine Translation Systems with Second Language Proficiency Tests
Takuya Matsuzaki, Akira Fujita, Naoya Todo and Noriko H. Arai 145

Representation Based Translation Evaluation Metrics
Boxing Chen and Hongyu Guo 150

Exploring the Planet of the APEs: a Comparative Study of State-of-the-art Methods for MT Automatic Post-Editing
Rajen Chatterjee, Marion Weller, Matteo Negri and Marco Turchi 156

Efficient Learning for Undirected Topic Models
Jiatao Gu and Victor O.K. Li 162

A Hassle-Free Unsupervised Domain Adaptation Method Using Instance Similarity Features
Jianfei Yu and Jing Jiang 168

Dependency-based Convolutional Neural Networks for Sentence Embedding
Mingbo Ma, Liang Huang, Bowen Zhou and Bing Xiang 174

Non-Linear Text Regression with a Deep Convolutional Neural Network
Zsolt Bitvai and Trevor Cohn 180

A Unified Learning Framework of Skip-Grams and Global Vectors
Jun Suzuki and Masaaki Nagata 186

Pre-training of Hidden-Unit CRFs
Young-Bum Kim, Karl Stratos and Ruhi Sarikaya 192

Distributional Neural Networks for Automatic Resolution of Crossword Puzzles
Aliaksei Severyn, Massimo Nicosia, Gianni Barlacchi and Alessandro Moschitti 199

Word Order Typology through Multilingual Word Alignment
Robert Östling 205

Measuring idiosyncratic interests in children with autism
Masoud Rouhizadeh, Emily Prud'hommeaux, Jan van Santen and Richard Sproat 212

Frame-Semantic Role Labeling with Heterogeneous Annotations
Meghana Kshirsagar, Sam Thomson, Nathan Schneider, Jaime Carbonell, Noah A. Smith and Chris Dyer 218

Semantic Interpretation of Superlative Expressions via Structured Knowledge Bases
Sheng Zhang, Yansong Feng, Songfang Huang, Kun Xu, Zhe Han and Dongyan Zhao 225

Grounding Semantics in Olfactory Perception
Douwe Kiela, Luana Bulat and Stephen Clark 231

Word-based Japanese typed dependency parsing with grammatical function analysis
Takaaki Tanaka and Masaaki Nagata 237

KLcpos3 - a Language Similarity Measure for Delexicalized Parser Transfer
Rudolf Rosa and Zdenek Zabokrtsky 243

CCG Supertagging with a Recurrent Neural Network
Wenduan Xu, Michael Auli and Stephen Clark 250

An Efficient Dynamic Oracle for Unrestricted Non-Projective Parsing
Carlos Gómez-Rodríguez and Daniel Fernández-González 256

Synthetic Word Parsing Improves Chinese Word Segmentation
Fei Cheng, Kevin Duh and Yuji Matsumoto 262

If all you have is a bit of the Bible: Learning POS taggers for truly low-resource languages
Željko Agić, Dirk Hovy and Anders Søgaard 268

Improving distant supervision using inference learning
Roland Roller, Eneko Agirre, Aitor Soroa and Mark Stevenson 273

A Lexicalized Tree Kernel for Open Information Extraction
Ying Xu, Christoph Ringlstetter, Mi-young Kim, Grzegorz Kondrak, Randy Goebel and Yusuke Miyao 279

A Dependency-Based Neural Network for Relation Classification
Yang Liu, Furu Wei, Sujian Li, Heng Ji, Ming Zhou and Houfeng WANG 285

Embedding Methods for Fine Grained Entity Type Classification
Dani Yogatama, Daniel Gillick and Nevena Lazic 291

Sieve-Based Entity Linking for the Biomedical Domain
Jennifer D'Souza and Vincent Ng 297

Open IE as an Intermediate Structure for Semantic Tasks
Gabriel Stanovsky and Ido Dagan 303

Recovering dropped pronouns from Chinese text messages
Yaqin Yang, Yalin Liu and Nianwen Xue 309

The Users Who Say 'Ni': Audience Identification in Chinese-language Restaurant Reviews
Rob Voigt and Dan Jurafsky 314

Chinese Zero Pronoun Resolution: A Joint Unsupervised Discourse-Aware Model Rivaling State-of-the-Art Resolvers
Chen Chen and Vincent Ng 320

Co-Simmate: Quick Retrieving All Pairwise Co-Simrank Scores
Yu Weiren and Julie McCann 327

Retrieval of Research-level Mathematical Information Needs: A Test Collection and Technical Terminology Experiment
Yiannos Stathopoulos and Simone Teufel 334

Learning to Mine Query Subtopics from Query Log
Zhenzhong Zhang, Le Sun and Xianpei Han 341

Learning Topic Hierarchies for Wikipedia Categories
Linmei Hu, Xuzhong Wang, Mengdi Zhang, Juanzi Li, Xiaoli Li, Chao Shao, Jie Tang and Yongbin Liu 346

Semantic Clustering and Convolutional Neural Network for Short Text Categorization
Peng Wang, Jiaming Xu, Bo Xu, Chenglin Liu, Heng Zhang, Fangyuan Wang and Hongwei Hao 352

Document Level Time-anchoring for TimeLine Extraction
Egoitz Laparra, Itziar Aldabe and German Rigau 358

Event Detection and Domain Adaptation with Convolutional Neural Networks
Thien Huu Nguyen and Ralph Grishman 365

Seed-Based Event Trigger Labeling: How far can event descriptions get us?
Ofer Bronstein, Ido Dagan, Qi Li, Heng Ji and Anette Frank 372

An Empirical Study of Chinese Name Matching and Applications
Nanyun Peng, Mo Yu and Mark Dredze 377

Language Identification and Modeling in Specialized Hardware
Kenneth Heafield, Rohan Kshirsagar and Santiago Barona 384

Cross-lingual Transfer of Named Entity Recognizers without Parallel Corpora
Ayah Zirikly 390

Robust Multi-Relational Clustering via ℓ_1-Norm Symmetric Nonnegative Matrix Factorization
Kai Liu and Hua Wang 397

Painless Labeling with Application to Text Mining
Sajib Dasgupta 402

FrameNet+: Fast Paraphrastic Tripling of FrameNet
Ellie Pavlick, Travis Wolfe, Pushpendre Rastogi, Chris Callison-Burch, Mark Dredze and Benjamin Van Durme 408

IWNLP: Inverse Wiktionary for Natural Language Processing
Matthias Liebeck and Stefan Conrad 414

TR9856: A Multi-word Term Relatedness Benchmark
Ran Levy, Liat Ein-Dor, Shay Hummel, Ruty Rinott and Noam Slonim 419

PPDB 2.0: Better paraphrase ranking, fine-grained entailment relations, word embeddings, and style classification
Ellie Pavlick, Pushpendre Rastogi, Juri Ganitkevitch, Benjamin Van Durme and Chris Callison-Burch 425

Automatic Discrimination between Cognates and Borrowings
Alina Maria Ciobanu and Liviu P. Dinu 431

The Media Frames Corpus: Annotations of Frames Across Issues
Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik and Noah A. Smith 438

deltaBLEU: A Discriminative Metric for Generation Tasks with Intrinsically Diverse Targets
Michel Galley, Chris Brockett, Alessandro Sordoni, Yangfeng Ji, Michael Auli, Chris Quirk, Margaret Mitchell, Jianfeng Gao and Bill Dolan 445

Tibetan Unknown Word Identification from News Corpora for Supporting Lexicon-based Tibetan Word Segmentation
Minghua Nuo, Huidan Liu, Congjun Long and Jian Wu 451

Learning Lexical Embeddings with Syntactic and Lexicographic Knowledge
Tong Wang, Abdelrahman Mohamed and Graeme Hirst 458

Non-distributional Word Vector Representations
Manaal Faruqui and Chris Dyer 464

Early and Late Combinations of Criteria for Reranking Distributional Thesauri
Olivier Ferret 470

Dependency length minimisation effects in short spans: a large-scale analysis of adjective placement in complex noun phrases
Kristina Gulordava, Paola Merlo and Benoit Crabbé 477

Tagging Performance Correlates with Author Age
Dirk Hovy and Anders Søgaard 483

User Based Aggregation for Biterm Topic Model
Weizheng Chen, Jinpeng Wang, Yan Zhang, Hongfei Yan and Xiaoming Li 489

The Fixed-Size Ordinally-Forgetting Encoding Method for Neural Network Language Models
ShiLiang Zhang, Hui Jiang, MingBin Xu, JunFeng Hou and LiRong Dai 495

Unsupervised Decomposition of a Multi-Author Document Based on Naive-Bayesian Model
Khaled Aldebei, Xiangjian He and Jie Yang 501

Extended Topic Model for Word Dependency
Tong Wang, Vish Viswanath and Ping Chen 506

Dependency Recurrent Neural Language Models for Sentence Completion
Piotr Mirowski and Andreas Vlachos 511

Point Process Modelling of Rumour Dynamics in Social Media
Michal Lukasik, Trevor Cohn and Kalina Bontcheva 518

Learning Hidden Markov Models with Distributed State Representations for Domain Adaptation
Min Xiao and Yuhong Guo 524

MT Quality Estimation for Computer-assisted Translation: Does it Really Help?
Marco Turchi, Matteo Negri and Marcello Federico 530

Context-Dependent Translation Selection Using Convolutional Neural Network
Baotian Hu, Zhaopeng Tu, Zhengdong Lu, Hang Li and Qingcai Chen 536

Learning Word Reorderings for Hierarchical Phrase-based Statistical Machine Translation
jingyi zhang, Masao Utiyama, Eiichro Sumita and Hai Zhao 542

UNRAVEL—A Decipherment Toolkit
Malte Nuhn, Julian Schamper and Hermann Ney 549

Multi-Pass Decoding With Complex Feature Guidance for Statistical Machine Translation
Benjamin Marie and Aurélien Max 554

What's in a Domain? Analyzing Genre and Topic Differences in Statistical Machine Translation
Marlies van der Wees, Arianna Bisazza, Wouter Weerkamp and Christof Monz 560

Learning Cross-lingual Word Embeddings via Matrix Co-factorization
Tianze Shi, Zhiyuan Liu, Yang Liu and Maosong Sun 567

Improving Pivot Translation by Remembering the Pivot
Akiva Miura, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura 573

BrailleSUM: A News Summarization System for the Blind and Visually Impaired People
Xiaojun Wan and Yue Hu 578

Automatic Identification of Age-Appropriate Ratings of Song Lyrics
Anggi Maulidyani and Ruli Manurung 583

Ground Truth for Grammaticality Correction Metrics
Courtney Napoles, Keisuke Sakaguchi, Matt Post and Joel Tetreault 588

Radical Embedding: Delving Deeper to Chinese Radicals
Xinlei Shi, Junjie Zhai, Xudong Yang, Zehua Xie and Chao Liu 594

Automatic Detection of Sentence Fragments
Chak Yan Yeung and John Lee 599

A Computational Approach to Automatic Prediction of Drunk-Texting
Aditya Joshi, Abhijit Mishra, Balamurali AR, Pushpak Bhattacharyya and Mark J Carman 604

Reducing infrequent-token perplexity via variational corpora
Yusheng Xie, Pranjal Daga, Yu Cheng, Kunpeng Zhang, Ankit Agrawal and Alok Choudhary . 609

A Hierarchical Knowledge Representation for Expert Finding on Social Media
Yanran Li, Wenjie Li and Sujian Li 616

Tackling Sparsity, the Achilles Heel of Social Networks: Language Model Smoothing via Social Regularization
Rui Yan, Xiang Li, Mengwen Liu and Xiaohua Hu 623

Twitter User Geolocation Using a Unified Text and Network Prediction Model
Afshin Rahimi, Trevor Cohn and Timothy Baldwin 630

Automatic Keyword Extraction on Twitter
Luis Marujo, Wang Ling, Isabel Trancoso, Chris Dyer, Alan W Black, Anatole Gershman, David Martins de Matos, João Neto and Jaime Carbonell 637

Towards a Contextual Pragmatic Model to Detect Irony in Tweets
Jihen Karoui, Benamara Farah, Véronique MORICEAU, Nathalie Aussenac-Gilles and Lamia Hadrich-Belguith 644

Annotation and Classification of an Email Importance Corpus
Fan Zhang and Kui Xu 651

Lexical Comparison Between Wikipedia and Twitter Corpora by Using Word Embeddings
Luchen Tan, Haotian Zhang, Charles Clarke and Mark Smucker 657

The Discovery of Natural Typing Annotations: User-produced Potential Chinese Word Delimiters
Dakui Zhang, Yu Mao, Yang Liu, Hanshi Wang, Chuyuan Wei and Shiping Tang 662

One Tense per Scene: Predicting Tense in Chinese Conversations
Tao Ge, Heng Ji, Baobao Chang and Zhifang Sui 668

A Language-Independent Feature Schema for Inflectional Morphology
John Sylak-Glassman, Christo Kirov, David Yarowsky and Roger Que 674

Rhetoric Map of an Answer to Compound Queries
Boris Galitsky, Dmitry Ilvovsky and Sergey O. Kuznetsov 681

Thread-Level Information for Comment Classification in Community Question Answering
Alberto Barrón-Cedeño, Simone Filice, Giovanni Da San Martino, Shafiq Joty, Lluís Màrquez, Preslav Nakov and Alessandro Moschitti 687

Learning Hybrid Representations to Retrieve Semantically Equivalent Questions
Cicero dos Santos, Luciano Barbosa, Dasha Bogdanova and Bianca Zadrozny 694

Machine Comprehension with Syntax, Frames, and Semantics
Hai Wang, Mohit Bansal, Kevin Gimpel and David McAllester 700

A Long Short-Term Memory Model for Answer Sentence Selection in Question Answering
Di Wang and Eric Nyberg 707

Answer Sequence Learning with Neural Networks for Answer Selection in Community Question Answering
Xiaoqiang Zhou, Baotian Hu, Qingcai Chen, Buzhou Tang and Xiaolong Wang 713

Bilingual Word Embeddings from Non-Parallel Document-Aligned Data Applied to Bilingual Lexicon Induction
Ivan Vulić and Marie-Francine Moens 719

How Well Do Distributional Models Capture Different Types of Semantic Knowledge?
Dana Rubinstein, Effi Levi, Roy Schwartz and Ari Rappoport 726

Low-Rank Tensors for Verbs in Compositional Distributional Semantics
Daniel Fried, Tamara Polajnar and Stephen Clark 731

Constrained Semantic Forests for Improved Discriminative Semantic Parsing
Wei Lu 737

Automatic Identification of Rhetorical Questions
Shohini Bhattasali, Jeremy Cytryn, Elana Feldman and Joonsuk Park 743

Lifelong Learning for Sentiment Classification
Zhiyuan Chen, Nianzu Ma and Bing Liu 750

Harnessing Context Incongruity for Sarcasm Detection
Aditya Joshi, Vinita Sharma and Pushpak Bhattacharyya 757

Emotion Detection in Code-switching Texts via Bilingual and Sentimental Information
Zhongqing Wang, Sophia Lee, Shoushan Li and Guodong Zhou 763

Model Adaptation for Personalized Opinion Analysis
Mohammad Al Boni, Keira Zhou, Hongning Wang and Matthew S. Gerber 769

Linguistic Template Extraction for Recognizing Reader-Emotion and Emotional Resonance Writing Assistance
Yung-Chun Chang, Cen-Chieh Chen, Yu-lun Hsieh, Chien Chin Chen and Wen-Lian Hsu 775

Aspect-Level Cross-lingual Sentiment Classification with Constrained SMT
Patrik Lambert 781

Predicting Valence-Arousal Ratings of Words Using a Weighted Graph Method
Liang-Chih Yu, Jin Wang, K. Robert Lai and Xue-jie Zhang 788

Multi-domain Dialog State Tracking using Recurrent Neural Networks
Nikola Mrkšić, Diarmuid Ó Séaghdha, Blaise Thomson, Milica Gasic, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen and Steve Young 794

Dialogue Management based on Sentence Clustering
Wendong Ge and Bo Xu 800

Compact Lexicon Selection with Spectral Methods
Young-Bum Kim, Karl Stratos, Xiaohu Liu and Ruhi Sarikaya 806

The Impact of Listener Gaze on Predicting Reference Resolution
Nikolina Koleva, Martin Villalba, Maria Staudte and Alexander Koller 812

A Simultaneous Recognition Framework for the Spoken Language Understanding Module of Intelligent Personal Assistant Software on Smart Phones
Changsu Lee, Youngjoong Ko and Jungyun Seo 818

A Deeper Exploration of the Standard PB-SMT Approach to Text Simplification and its Evaluation
Sanja Štajner, Hannah Bechara and Horacio Saggion 823

Learning Summary Prior Representation for Extractive Summarization
Ziqiang Cao, Furu Wei, Sujian Li, Wenjie Li, Ming Zhou and Houfeng WANG 829

A Methodology for Evaluating Timeline Generation Algorithms based on Deep Semantic Units
Sandro Bauer and Simone Teufel 834

Unsupervised extractive summarization via coverage maximization with syntactic and semantic concepts
Natalie Schluter and Anders Søgaard 840

Low Resource Dependency Parsing: Cross-lingual Parameter Sharing in a Neural Network Parser
Long Duong, Trevor Cohn, Steven Bird and Paul Cook 845

Semantic Structure Analysis of Noun Phrases using Abstract Meaning Representation
Yuichiro Sawai, Hiroyuki Shindo and Yuji Matsumoto 851

Boosting Transition-based AMR Parsing with Refined Actions and Auxiliary Analyzers
Chuan Wang, Nianwen Xue and Sameer Pradhan 857

Generative Incremental Dependency Parsing with Neural Networks
Jan Buys and Phil Blunsom 863

Labeled Grammar Induction with Minimal Supervision
Yonatan Bisk, Christos Christodoulopoulos and Julia Hockenmaier 870

On the Importance of Ezafe Construction in Persian Parsing
Alireza Nourian, Mohammad Sadegh Rasooli, Mohsen Imany and Heshaam Faili 877

Conference Program

Monday, July 27

17:00–18:00 Session 4: Short Papers

Session 4A: 17:00–18:00 Semantics

A Framework for the Construction of Monolingual and Cross-lingual Word Similarity Datasets
José Camacho-Collados, Mohammad Taher Pilehvar and Roberto Navigli

On metric embedding for boosting semantic similarity computations
Julien Subercaze, Christophe Gravier and Frédérique Laforest

Improving Distributed Representation of Word Sense via WordNet Gloss Composition and Context Clustering
Tao Chen, Ruifeng Xu, Yulan He and Xuan Wang

A Multitask Objective to Inject Lexical Contrast into Distributional Semantics
Nghia The Pham, Angeliki Lazaridou and Marco Baroni

Session 4B: 17:00–18:00 Sentiment Analysis

Semi-Stacking for Semi-supervised Sentiment Classification
Shoushan Li, Lei Huang, Jingjing Wang and Guodong Zhou

Deep Markov Neural Network for Sequential Data Classification
Min Yang

Semantic Analysis and Helpfulness Prediction of Text for Online Product Reviews
Yinfei Yang, Yaowei Yan, Minghui Qiu and Forrest Bao

Document Classification by Inversion of Distributed Language Representations
Matt Taddy

Monday, July 27 (continued)

Session 4C: 17:00–18:00 Summarization and Generation

Using Tweets to Help Sentence Compression for News Highlights Generation
Zhongyu Wei, Yang Liu, Chen Li and Wei Gao

Domain-Specific Paraphrase Extraction
Ellie Pavlick, Juri Ganitkevitch, Tsz Ping Chan, Xuchen Yao, Benjamin Van Durme and Chris Callison-Burch

Simplifying Lexical Simplification: Do We Need Simplified Corpora?
Goran Glavaš and Sanja Štajner

Zoom: a corpus of natural language descriptions of map locations
Romina Altamirano, Thiago Ferreira, Ivandré Paraboni and Luciana Benotti

Session 4D: 17:00–18:00 Discourse, Coreference

Generating overspecified referring expressions: the role of discrimination
Ivandré Paraboni, Michelle Galindo and Douglas Iacovelli

Using prosodic annotations to improve coreference resolution of spoken text
Ina Roesiger and Arndt Riester

Spectral Semi-Supervised Discourse Relation Classification
Robert Fisher and Reid Simmons

I do not disagree: leveraging monolingual alignment to detect disagreement in dialogue
Ajda Gokcen and Marie-Catherine de Marneffe

Monday, July 27 (continued)

Session 4E: 17:00–18:00 Language and Vision

Language Models for Image Captioning: The Quirks and What Works
Jacob Devlin, Hao Cheng, Hao Fang, Saurabh Gupta, Li Deng, Xiaodong He, Geoffrey Zweig and Margaret Mitchell

A Distributed Representation Based Query Expansion Approach for Image Captioning
Semih Yagcioglu, Erkut Erdem, Aykut Erdem and Ruket Cakici

Learning language through pictures
Grzegorz Chrupała, Ákos Kádár and Afra Alishahi

Exploiting Image Generality for Lexical Entailment Detection
Douwe Kiela, Laura Rimell, Ivan Vulić and Stephen Clark

Tuesday, July 28

10:30–12:00 Session 5: Short Papers

Session 5A: 10:30–12:00 Machine Translation

Lexicon Stratification for Translating Out-of-Vocabulary Words
Yulia Tsvetkov and Chris Dyer

Recurrent Neural Network based Rule Sequence Model for Statistical Machine Translation
Heng Yu and Xuan Zhu

Discriminative Preordering Meets Kendall's τ Maximization
Sho Hoshino, Yusuke Miyao, Katsuhito Sudoh, Katsuhiko Hayashi and Masaaki Nagata

Evaluating Machine Translation Systems with Second Language Proficiency Tests
Takuya Matsuzaki, Akira Fujita, Naoya Todo and Noriko H. Arai

Representation Based Translation Evaluation Metrics
Boxing Chen and Hongyu Guo

Exploring the Planet of the APEs: a Comparative Study of State-of-the-art Methods for MT Automatic Post-Editing
Rajen Chatterjee, Marion Weller, Matteo Negri and Marco Turchi

Session 5B: 10:30–12:00 Machine Learning and Topic Modeling

Efficient Learning for Undirected Topic Models
Jiatao Gu and Victor O.K. Li

A Hassle-Free Unsupervised Domain Adaptation Method Using Instance Similarity Features
Jianfei Yu and Jing Jiang

Dependency-based Convolutional Neural Networks for Sentence Embedding
Mingbo Ma, Liang Huang, Bowen Zhou and Bing Xiang

Non-Linear Text Regression with a Deep Convolutional Neural Network
Zsolt Bitvai and Trevor Cohn

A Unified Learning Framework of Skip-Grams and Global Vectors
Jun Suzuki and Masaaki Nagata

Pre-training of Hidden-Unit CRFs
Young-Bum Kim, Karl Stratos and Ruhi Sarikaya

Session 5C: 10:30–12:00 Semantics, Linguistic and Psycholinguistic Aspects of CL

Distributional Neural Networks for Automatic Resolution of Crossword Puzzles
Aliaksei Severyn, Massimo Nicosia, Gianni Barlacchi and Alessandro Moschitti

Word Order Typology through Multilingual Word Alignment
Robert Östling

Measuring idiosyncratic interests in children with autism
Masoud Rouhizadeh, Emily Prud'hommeaux, Jan van Santen and Richard Sproat

Frame-Semantic Role Labeling with Heterogeneous Annotations
Meghana Kshirsagar, Sam Thomson, Nathan Schneider, Jaime Carbonell, Noah A. Smith and Chris Dyer

Semantic Interpretation of Superlative Expressions via Structured Knowledge Bases
Sheng Zhang, Yansong Feng, Songfang Huang, Kun Xu, Zhe Han and Dongyan Zhao

Grounding Semantics in Olfactory Perception
Douwe Kiela, Luana Bulat and Stephen Clark

Session 5D: 10:30–12:00 Parsing, Tagging

Word-based Japanese typed dependency parsing with grammatical function analysis
Takaaki Tanaka and Masaaki Nagata

KLcpos3 - a Language Similarity Measure for Delexicalized Parser Transfer
Rudolf Rosa and Zdenek Zabokrtsky

CCG Supertagging with a Recurrent Neural Network
Wenduan Xu, Michael Auli and Stephen Clark

An Efficient Dynamic Oracle for Unrestricted Non-Projective Parsing
Carlos Gómez-Rodríguez and Daniel Fernández-González

Synthetic Word Parsing Improves Chinese Word Segmentation
Fei Cheng, Kevin Duh and Yuji Matsumoto

If all you have is a bit of the Bible: Learning POS taggers for truly low-resource languages
Željko Agić, Dirk Hovy and Anders Søgaard

Session 5E: 10:30–12:00 Information Extraction

Improving distant supervision using inference learning
Roland Roller, Eneko Agirre, Aitor Soroa and Mark Stevenson

A Lexicalized Tree Kernel for Open Information Extraction
Ying Xu, Christoph Ringlstetter, Mi-young Kim, Grzegorz Kondrak, Randy Goebel and Yusuke Miyao

A Dependency-Based Neural Network for Relation Classification
Yang Liu, Furu Wei, Sujian Li, Heng Ji, Ming Zhou and Houfeng WANG

Embedding Methods for Fine Grained Entity Type Classification
Dani Yogatama, Daniel Gillick and Nevena Lazic

Sieve-Based Entity Linking for the Biomedical Domain
Jennifer D'Souza and Vincent Ng

Open IE as an Intermediate Structure for Semantic Tasks
Gabriel Stanovsky and Ido Dagan

16:30–19:30 Poster and Dinner Session 2: Short Papers, Student Research Workshop Papers

Session P2.01: 16:30–19:30 Poster: Discourse, Pragmatics

Recovering dropped pronouns from Chinese text messages
Yaqin Yang, Yalin Liu and Nianwen Xue

The Users Who Say 'Ni': Audience Identification in Chinese-language Restaurant Reviews
Rob Voigt and Dan Jurafsky

Chinese Zero Pronoun Resolution: A Joint Unsupervised Discourse-Aware Model Rivaling State-of-the-Art Resolvers
Chen Chen and Vincent Ng

Session P2.02: 16:30–19:30 Poster: Information Retrieval

Co-Simmate: Quick Retrieving All Pairwise Co-Simrank Scores
Yu Weiren and Julie McCann

Retrieval of Research-level Mathematical Information Needs: A Test Collection and Technical Terminology Experiment
Yiannos Stathopoulos and Simone Teufel

Learning to Mine Query Subtopics from Query Log
Zhenzhong Zhang, Le Sun and Xianpei Han

Session P2.03: 16:30–19:30 Poster: Information Extraction and Text Mining

Learning Topic Hierarchies for Wikipedia Categories
Linmei Hu, Xuzhong Wang, Mengdi Zhang, Juanzi Li, Xiaoli Li, Chao Shao, Jie Tang and Yongbin Liu

Semantic Clustering and Convolutional Neural Network for Short Text Categorization
Peng Wang, Jiaming Xu, Bo Xu, Chenglin Liu, Heng Zhang, Fangyuan Wang and Hongwei Hao

Document Level Time-anchoring for TimeLine Extraction
Egoitz Laparra, Itziar Aldabe and German Rigau

Event Detection and Domain Adaptation with Convolutional Neural Networks
Thien Huu Nguyen and Ralph Grishman

Seed-Based Event Trigger Labeling: How far can event descriptions get us?
Ofer Bronstein, Ido Dagan, Qi Li, Heng Ji and Anette Frank

An Empirical Study of Chinese Name Matching and Applications
Nanyun Peng, Mo Yu and Mark Dredze

Language Identification and Modeling in Specialized Hardware
Kenneth Heafield, Rohan Kshirsagar and Santiago Barona

Cross-lingual Transfer of Named Entity Recognizers without Parallel Corpora
Ayah Zirikly

Robust Multi-Relational Clustering via ℓ_1-Norm Symmetric Nonnegative Matrix Factorization
Kai Liu and Hua Wang

Session P2.05: 16:30–19:30 Poster: Language Resources

Painless Labeling with Application to Text Mining
Sajib Dasgupta

FrameNet+: Fast Paraphrastic Tripling of FrameNet
Ellie Pavlick, Travis Wolfe, Pushpendre Rastogi, Chris Callison-Burch, Mark Dredze and Benjamin Van Durme

IWNLP: Inverse Wiktionary for Natural Language Processing
Matthias Liebeck and Stefan Conrad

TR9856: A Multi-word Term Relatedness Benchmark
Ran Levy, Liat Ein-Dor, Shay Hummel, Ruty Rinott and Noam Slonim

PPDB 2.0: Better paraphrase ranking, fine-grained entailment relations, word embeddings, and style classification
Ellie Pavlick, Pushpendre Rastogi, Juri Ganitkevitch, Benjamin Van Durme and Chris Callison-Burch

Automatic Discrimination between Cognates and Borrowings
Alina Maria Ciobanu and Liviu P. Dinu

The Media Frames Corpus: Annotations of Frames Across Issues
Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik and Noah A. Smith

deltaBLEU: A Discriminative Metric for Generation Tasks with Intrinsically Diverse Targets
Michel Galley, Chris Brockett, Alessandro Sordoni, Yangfeng Ji, Michael Auli, Chris Quirk, Margaret Mitchell, Jianfeng Gao and Bill Dolan

Tibetan Unknown Word Identification from News Corpora for Supporting Lexicon-based Tibetan Word Segmentation
Minghua Nuo, Huidan Liu, Congjun Long and Jian Wu

Session P2.06: 16:30–19:30 Poster: Lexical Semantics

Learning Lexical Embeddings with Syntactic and Lexicographic Knowledge
Tong Wang, Abdelrahman Mohamed and Graeme Hirst

Non-distributional Word Vector Representations
Manaal Faruqui and Chris Dyer

Early and Late Combinations of Criteria for Reranking Distributional Thesauri
Olivier Ferret

Session P2.07: 16:30–19:30 Poster: Linguistic and Psycholinguistic Aspects of CL

Dependency length minimisation effects in short spans: a large-scale analysis of adjective placement in complex noun phrases
Kristina Gulordava, Paola Merlo and Benoit Crabbé

Tagging Performance Correlates with Author Age
Dirk Hovy and Anders Søgaard

Session P2.08: 16:30–19:30 Poster: Machine Learning and Topic Models

User Based Aggregation for Biterm Topic Model
Weizheng Chen, Jinpeng Wang, Yan Zhang, Hongfei Yan and Xiaoming Li

The Fixed-Size Ordinally-Forgetting Encoding Method for Neural Network Language Models
ShiLiang Zhang, Hui Jiang, MingBin Xu, JunFeng Hou and LiRong Dai

Unsupervised Decomposition of a Multi-Author Document Based on Naive-Bayesian Model
Khaled Aldebei, Xiangjian He and Jie Yang

Extended Topic Model for Word Dependency
Tong Wang, Vish Viswanath and Ping Chen

Dependency Recurrent Neural Language Models for Sentence Completion
Piotr Mirowski and Andreas Vlachos

Point Process Modelling of Rumour Dynamics in Social Media
Michal Lukasik, Trevor Cohn and Kalina Bontcheva

Learning Hidden Markov Models with Distributed State Representations for Domain Adaptation
Min Xiao and Yuhong Guo

Session P2.09: 16:30–19:30 Poster: Machine Translation

MT Quality Estimation for Computer-assisted Translation: Does it Really Help?
Marco Turchi, Matteo Negri and Marcello Federico

Context-Dependent Translation Selection Using Convolutional Neural Network
Baotian Hu, Zhaopeng Tu, Zhengdong Lu, Hang Li and Qingcai Chen

Learning Word Reorderings for Hierarchical Phrase-based Statistical Machine Translation
jingyi zhang, Masao Utiyama, Eiichro Sumita and Hai Zhao

UNRAVEL—A Decipherment Toolkit
Malte Nuhn, Julian Schamper and Hermann Ney

Multi-Pass Decoding With Complex Feature Guidance for Statistical Machine Translation
Benjamin Marie and Aurélien Max

What's in a Domain? Analyzing Genre and Topic Differences in Statistical Machine Translation
Marlies van der Wees, Arianna Bisazza, Wouter Weerkamp and Christof Monz

Learning Cross-lingual Word Embeddings via Matrix Co-factorization
Tianze Shi, Zhiyuan Liu, Yang Liu and Maosong Sun

Session P2.10: 16:30–19:30 Poster: NLP Applications

Improving Pivot Translation by Remembering the Pivot
Akiva Miura, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura

BrailleSUM: A News Summarization System for the Blind and Visually Impaired People
Xiaojun Wan and Yue Hu

Automatic Identification of Age-Appropriate Ratings of Song Lyrics
Anggi Maulidyani and Ruli Manurung

Ground Truth for Grammaticality Correction Metrics
Courtney Napoles, Keisuke Sakaguchi, Matt Post and Joel Tetreault

Radical Embedding: Delving Deeper to Chinese Radicals
Xinlei Shi, Junjie Zhai, Xudong Yang, Zehua Xie and Chao Liu

Automatic Detection of Sentence Fragments
Chak Yan Yeung and John Lee

Session P2.11: 16:30–19:30 Poster: NLP for the Web and Social Media

A Computational Approach to Automatic Prediction of Drunk-Texting
Aditya Joshi, Abhijit Mishra, Balamurali AR, Pushpak Bhattacharyya and Mark J Carman

Reducing infrequent-token perplexity via variational corpora
Yusheng Xie, Pranjal Daga, Yu Cheng, Kunpeng Zhang, Ankit Agrawal and Alok Choudhary

A Hierarchical Knowledge Representation for Expert Finding on Social Media
Yanran Li, Wenjie Li and Sujian Li

Tackling Sparsity, the Achilles Heel of Social Networks: Language Model Smoothing via Social Regularization
Rui Yan, Xiang Li, Mengwen Liu and Xiaohua Hu

Twitter User Geolocation Using a Unified Text and Network Prediction Model
Afshin Rahimi, Trevor Cohn and Timothy Baldwin

Automatic Keyword Extraction on Twitter
Luis Marujo, Wang Ling, Isabel Trancoso, Chris Dyer, Alan W Black, Anatole Gershman, David Martins de Matos, João Neto and Jaime Carbonell

Towards a Contextual Pragmatic Model to Detect Irony in Tweets
Jihen Karoui, Benamara Farah, Véronique MORICEAU, Nathalie Aussenac-Gilles and Lamia Hadrich-Belguith

Annotation and Classification of an Email Importance Corpus
Fan Zhang and Kui Xu

Lexical Comparison Between Wikipedia and Twitter Corpora by Using Word Embeddings
Luchen Tan, Haotian Zhang, Charles Clarke and Mark Smucker

Session P2.12: 16:30–19:30 Poster: Morphology, Word Segmentation

The Discovery of Natural Typing Annotations: User-produced Potential Chinese Word Delimiters
Dakui Zhang, Yu Mao, Yang Liu, Hanshi Wang, Chuyuan Wei and Shiping Tang

One Tense per Scene: Predicting Tense in Chinese Conversations
Tao Ge, Heng Ji, Baobao Chang and Zhifang Sui

A Language-Independent Feature Schema for Inflectional Morphology
John Sylak-Glassman, Christo Kirov, David Yarowsky and Roger Que

Session P2.13: 16:30–19:30 Poster: Question Answering

Rhetoric Map of an Answer to Compound Queries
Boris Galitsky, Dmitry Ilvovsky and Sergey O. Kuznetsov

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

Learning Hybrid Representations to Retrieve Semantically Equivalent Questions
Cicero dos Santos, Luciano Barbosa, Dasha Bogdanova and Bianca Zadrozny

Machine Comprehension with Syntax, Frames, and Semantics
Hai Wang, Mohit Bansal, Kevin Gimpel and David McAllester

A Long Short-Term Memory Model for Answer Sentence Selection in Question Answering
Di Wang and Eric Nyberg

Answer Sequence Learning with Neural Networks for Answer Selection in Community Question Answering
Xiaoqiang Zhou, Baotian Hu, Qingcai Chen, Buzhou Tang and Xiaolong Wang

Session P2.14: 16:30–19:30 Poster: Semantics

Bilingual Word Embeddings from Non-Parallel Document-Aligned Data Applied to Bilingual Lexicon Induction
Ivan Vulić and Marie-Francine Moens

How Well Do Distributional Models Capture Different Types of Semantic Knowledge?
Dana Rubinstein, Effi Levi, Roy Schwartz and Ari Rappoport

Low-Rank Tensors for Verbs in Compositional Distributional Semantics
Daniel Fried, Tamara Polajnar and Stephen Clark

Constrained Semantic Forests for Improved Discriminative Semantic Parsing
Wei Lu

Automatic Identification of Rhetorical Questions
Shohini Bhattasali, Jeremy Cytryn, Elana Feldman and Joonsuk Park

Session P2.15: 16:30–19:30 Poster: Sentiment Analysis

Lifelong Learning for Sentiment Classification
Zhiyuan Chen, Nianzu Ma and Bing Liu

Harnessing Context Incongruity for Sarcasm Detection
Aditya Joshi, Vinita Sharma and Pushpak Bhattacharyya

Emotion Detection in Code-switching Texts via Bilingual and Sentimental Information
Zhongqing Wang, Sophia Lee, Shoushan Li and Guodong Zhou

Model Adaptation for Personalized Opinion Analysis
Mohammad Al Boni, Keira Zhou, Hongning Wang and Matthew S. Gerber

Linguistic Template Extraction for Recognizing Reader-Emotion and Emotional Resonance Writing Assistance
Yung-Chun Chang, Cen-Chieh Chen, Yu-lun Hsieh, Chien Chin Chen and Wen-Lian Hsu

Aspect-Level Cross-lingual Sentiment Classification with Constrained SMT
Patrik Lambert

Predicting Valence-Arousal Ratings of Words Using a Weighted Graph Method
Liang-Chih Yu, Jin Wang, K. Robert Lai and Xue-jie Zhang

Session P2.16: 16:30–19:30 Poster: Spoken Language Processing and Dialogue Systems

Multi-domain Dialog State Tracking using Recurrent Neural Networks
Nikola Mrkšić, Diarmuid Ó Séaghdha, Blaise Thomson, Milica Gasic, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen and Steve Young

Dialogue Management based on Sentence Clustering
Wendong Ge and Bo Xu

Compact Lexicon Selection with Spectral Methods
Young-Bum Kim, Karl Stratos, Xiaohu Liu and Ruhi Sarikaya

The Impact of Listener Gaze on Predicting Reference Resolution
Nikolina Koleva, Martin Villalba, Maria Staudte and Alexander Koller

A Simultaneous Recognition Framework for the Spoken Language Understanding Module of Intelligent Personal Assistant Software on Smart Phones
Changsu Lee, Youngjoong Ko and Jungyun Seo

Session P2.17: 16:30–19:30 Poster: Summarization and Generation

A Deeper Exploration of the Standard PB-SMT Approach to Text Simplification and its Evaluation
Sanja Štajner, Hannah Bechara and Horacio Saggion

Learning Summary Prior Representation for Extractive Summarization
Ziqiang Cao, Furu Wei, Sujian Li, Wenjie Li, Ming Zhou and Houfeng WANG

A Methodology for Evaluating Timeline Generation Algorithms based on Deep Semantic Units
Sandro Bauer and Simone Teufel

Unsupervised extractive summarization via coverage maximization with syntactic and semantic concepts
Natalie Schluter and Anders Søgaard

Session P2.18: 16:30–19:30 Poster: Parsing

Low Resource Dependency Parsing: Cross-lingual Parameter Sharing in a Neural Network Parser
Long Duong, Trevor Cohn, Steven Bird and Paul Cook

Semantic Structure Analysis of Noun Phrases using Abstract Meaning Representation
Yuichiro Sawai, Hiroyuki Shindo and Yuji Matsumoto

Boosting Transition-based AMR Parsing with Refined Actions and Auxiliary Analyzers
Chuan Wang, Nianwen Xue and Sameer Pradhan

Generative Incremental Dependency Parsing with Neural Networks
Jan Buys and Phil Blunsom

Labeled Grammar Induction with Minimal Supervision
Yonatan Bisk, Christos Christodoulopoulos and Julia Hockenmaier

On the Importance of Ezafe Construction in Persian Parsing
Alireza Nourian, Mohammad Sadegh Rasooli, Mohsen Imany and Heshaam Faili

Day Date

Session Ses Code: Ses Time–Ses End Time Ses Title

Gen *Time–Gen End Time Gen Title*
Gen Presenter