Conference Program

Thursday, July 30, 2015 – Rooms: 402A+402B

08:45-09:00 Opening Remarks

09:00-10:10 Session 1.a: Embedding input and output representations

Multichannel Variable-Size Convolution for Sentence Classification
Wenpeng Yin and Hinrich Schütze

Task-Oriented Learning of Word Embeddings for Semantic Relation Classification
Kazuma Hashimoto, Pontus Stenetorp, Makoto Miwa and Yoshimasa Tsuruoka

Symmetric Pattern Based Word Embeddings for Improved Word Similarity Prediction
Roy Schwartz, Roi Reichart and Ari Rappoport

A Coactive Learning View of Online Structured Prediction in Statistical Machine Translation
Artem Sokolov, Stefan Riezler and Shay B. Cohen

10:10-10:30 Session 1.b: Entity Linking (spotlight presentations)

A Joint Framework for Coreference Resolution and Mention Head Detection
Haoruo Peng, Kai-Wei Chang and Dan Roth

Entity Linking Korean Text: An Unsupervised Learning Approach using Semantic Relations
Youngsik Kim and Key-Sun Choi

Linking Entities Across Images and Text
Rebecka Weegar, Kalle Åström and Pierre Nugues

Recovering Traceability Links in Requirements Documents
Zeheng Li, Mingrui Chen, LiGuo Huang and Vincent Ng

10:30-11:00 Coffee Break

11:00-12:00 Session 2.a: Keynote Talk

On Spectral Graphical Models, and a New Look at Latent Variable Modeling in Natural Language Processing
Eric Xing, Carnegie Mellon University

12:00-12:30 Session 2.b: Short Paper Spotlights

Deep Neural Language Models for Machine Translation
Thang Luong, Michael Kayser and Christopher D. Manning

Reading behavior predicts syntactic categories
Maria Barrett and Anders Søgaard

One Million Sense-Tagged Instances for Word Sense Disambiguation and Induction
Kaveh Taghipour and Hwee Tou Ng

Model Selection for Type-Supervised Learning with Application to POS Tagging
Kristina Toutanova, Waleed Ammar, Pallavi Choudhury and Hoifung Poon

Feature Selection for Short Text Classification using Wavelet Packet Transform
Anuj Mahajan, Sharmistha Jat and Shourya Roy

Do dependency parsing metrics correlate with human judgments?
Barbara Plank, Héctor Martínez Alonso, Željko Agić, Danijela Merkler and Anders Søgaard

Learning Adjective Meanings with a Tensor-Based Skip-Gram Model
Jean Maillard and Stephen Clark

Accurate Cross-lingual Projection between Count-based Word Vectors by Exploiting Translatable Context Pairs
Shonosuke Ishiwatari, Nobuhiro Kaji, Naoki Yoshinaga, Masashi Toyoda and Masaru Kitsuregawa

Finding Opinion Manipulation Trolls in News Community Forums
Todor Mihaylov, Georgi Georgiev and Preslav Nakov

12:30-14:00 Lunch Break

14:00-15:30 Session 3: CoNLL Shared Task

The CoNLL-2015 Shared Task on Shallow Discourse Parsing
Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Rashmi Prasad, Christopher Bryant, Attapol Rutherford

A Refined End-to-End Discourse Parser
Jianxiang Wang and Man Lan

The UniTN Discourse Parser in CoNLL 2015 Shared Task: Token-level Sequence Labeling with Argument-specific Models
Evgeny Stepanov, Giuseppe Riccardi and Ali Orkan Bayer

The SoNLP-DP System in the CoNLL-2015 shared Task
Fang Kong, Sheng Li and Guodong Zhou

15:30-16:00 Coffee Break

16:00-17:10 Session 4.a: Syntactic Parsing

A Supertag-Context Model for Weakly-Supervised CCG Parser Learning
Dan Garrette, Chris Dyer, Jason Baldridge and Noah A. Smith

Transition-based Spinal Parsing
Miguel Ballesteros and Xavier Carreras

Cross-lingual Transfer for Unsupervised Dependency Parsing Without Parallel Data
Long Duong, Trevor Cohn, Steven Bird and Paul Cook

Incremental Recurrent Neural Network Dependency Parser with Search-based Discriminative Training
Majid Yazdani and James Henderson

17:10-17:30 Session 4.b: Cross-language studies (spotlight presentations)

Structural and lexical factors in adjective placement in complex noun phrases across Romance languages
Kristina Gulordava and Paola Merlo

Instance Selection Improves Cross-Lingual Model Training for Fine-Grained Sentiment Analysis
Roman Klinger and Philipp Cimiano

Annotation Projection-based Representation Learning for Cross-lingual Dependency Parsing
Min Xiao and Yuhong Guo

Friday, July 31, 2015 – Rooms: 402A+402B

09:00-10:10 Session 5.a: Semantics

An Iterative Similarity based Adaptation Technique for Cross-domain Text Classification
Himanshu Sharad Bhatt, Deepali Semwal and Shourya Roy

Detecting Semantically Equivalent Questions in Online User Forums
Dasha Bogdanova, Cicero dos Santos, Luciano Barbosa and Bianca Zadrozny

Making the Most of Crowdsourced Document Annotations: Confused Supervised LDA
Paul Felt, Eric Ringger, Jordan Boyd-Graber and Kevin Seppi

Big Data Small Data, In Domain Out-of Domain, Known Word Unknown Word: The Impact of Word Representations on Sequence Labelling Tasks
Lizhen Qu, Gabriela Ferraro, Liyuan Zhou, Weiwei Hou, Nathan Schneider and Timothy Baldwin

10:10-10:30 Session 5.b: Extraction and Labeling (spotlight presentations)

Labeled Morphological Segmentation with Semi-Markov Models
Ryan Cotterell, Thomas Müller, Alexander Fraser and Hinrich Schütze

Opinion Holder and Target Extraction based on the Induction of Verbal Categories
Michael Wiegand and Josef Ruppenhofer

Temporal Information Extraction from Korean Texts
Young-Seob Jeong, Zae Myung Kim, Hyun-Woo Do, Chae-Gyun Lim and Ho-Jin Choi

10:30-11:00 Coffee Break

11:00-12:00 Session 6.a: Keynote Talk

Does the Success of Deep Neural Network Language Processing Mean – Finally! – the End of Theoretical Linguistics?
Paul Smolensky, Johns Hopkins University

12:00-12:30 Session 6.b: Business Meeting

12:30-14:00 Lunch Break

14:00-15:10 Session 7.a: Language Structure

Cross-lingual syntactic variation over age and gender
Anders Johannsen, Dirk Hovy and Anders Søgaard

A Synchronous Hyperedge Replacement Grammar based approach for AMR parsing
Xiaochang Peng, Linfeng Song and Daniel Gildea

Contrastive Analysis with Predictive Power: Typology Driven Estimation of Grammatical Error Distributions in ESL
Yevgeni Berzak, Roi Reichart and Boris Katz

AIDA2: A Hybrid Approach for Token and Sentence Level Dialect Identification in Arabic
Mohamed Al-Badrashiny, Heba Elfardy and Mona Diab

15:10-15:30 Session 7.b: CoNLL Mix (spotlight presentations)

Analyzing Optimization for Statistical Machine Translation: MERT Learns Verbosity, PRO Learns Length
Francisco Guzmán, Preslav Nakov and Stephan Vogel

Learning to Exploit Structured Resources for Lexical Inference
Vered Shwartz, Omer Levy, Ido Dagan and Jacob Goldberger

Quantity, Contrast, and Convention in Cross-Situated Language Comprehension
Ian Perera and James Allen

15:30-16:00 Coffee Break

16:00-17:30 Session 8.a: Joint Poster Presentation (long, short and shared task papers)

Long Papers

A Joint Framework for Coreference Resolution and Mention Head Detection
Haoruo Peng, Kai-Wei Chang and Dan Roth

Entity Linking Korean Text: An Unsupervised Learning Approach using Semantic Relations
Youngsik Kim and Key-Sun Choi

Linking Entities Across Images and Text
Rebecka Weegar, Kalle Åström and Pierre Nugues

Recovering Traceability Links in Requirements Documents
Zeheng Li, Mingrui Chen, LiGuo Huang and Vincent Ng

Structural and lexical factors in adjective placement in complex noun phrases across Romance languages
Kristina Gulordava and Paola Merlo

Instance Selection Improves Cross-Lingual Model Training for Fine-Grained Sentiment Analysis
Roman Klinger and Philipp Cimiano

Annotation Projection-based Representation Learning for Cross-lingual Dependency Parsing
Min Xiao and Yuhong Guo

Labeled Morphological Segmentation with Semi-Markov Models
Ryan Cotterell, Thomas Müller, Alexander Fraser and Hinrich Schütze

Opinion Holder and Target Extraction based on the Induction of Verbal Categories
Michael Wiegand and Josef Ruppenhofer

Temporal Information Extraction from Korean Texts
Young-Seob Jeong, Zae Myung Kim, Hyun-Woo Do, Chae-Gyun Lim and Ho-Jin Choi

Analyzing Optimization for Statistical Machine Translation: MERT Learns Verbosity, PRO Learns Length
Francisco Guzmán, Preslav Nakov and Stephan Vogel

Learning to Exploit Structured Resources for Lexical Inference
Vered Shwartz, Omer Levy, Ido Dagan and Jacob Goldberger

Quantity, Contrast, and Convention in Cross-Situated Language Comprehension
Ian Perera and James Allen

Short Papers

Deep Neural Language Models for Machine Translation
Thang Luong, Michael Kayser and Christopher D. Manning

Reading behavior predicts syntactic categories
Maria Barrett and Anders Søgaard

One Million Sense-Tagged Instances for Word Sense Disambiguation and Induction
Kaveh Taghipour and Hwee Tou Ng

Model Selection for Type-Supervised Learning with Application to POS Tagging
Kristina Toutanova, Waleed Ammar, Pallavi Choudhury and Hoifung Poon

Feature Selection for Short Text Classification using Wavelet Packet Transform
Anuj Mahajan, Sharmistha Jat and Shourya Roy

Do dependency parsing metrics correlate with human judgments?
Barbara Plank, Héctor Martínez Alonso, Željko Agić, Danijela Merkler and Anders Søgaard

Learning Adjective Meanings with a Tensor-Based Skip-Gram Model
Jean Maillard and Stephen Clark

Accurate Cross-lingual Projection between Count-based Word Vectors by Exploiting Translatable Context Pairs
Shonosuke Ishiwatari, Nobuhiro Kaji, Naoki Yoshinaga, Masashi Toyoda and Masaru Kitsuregawa

Finding Opinion Manipulation Trolls in News Community Forums
Todor Mihaylov, Georgi Georgiev and Preslav Nakov

Shared Task Papers

A Hybrid Discourse Relation Parser in CoNLL 2015
Sobha Lalitha Devi, Sindhuja Gopalan, Lakshmi S, Pattabhi RK Rao, Vijay Sundar Ram and Malarkodi C.S.

A Minimalist Approach to Shallow Discourse Parsing and Implicit Relation Recognition
Christian Chiarcos and Niko Schenk

A Shallow Discourse Parsing System Based On Maximum Entropy Model
Jia Sun, Peijia Li, Weiqun Xu and Yonghong Yan

Hybrid Approach to PDTB-styled Discourse Parsing for CoNLL-2015
Yasuhisa Yoshida, Katsuhiko Hayashi, Tsutomu Hirao and Masaaki Nagata

Improving a Pipeline Architecture for Shallow Discourse Parsing
Yangqiu Song, Haoruo Peng, Parisa Kordjamshidi, Mark Sammons and Dan Roth

JAIST: A two-phase machine learning approach for identifying discourse relations in newswire texts
Son Nguyen, Quoc Ho and Minh Nguyen

Shallow Discourse Parsing Using Constituent Parsing Tree
Changge Chen, Peilu Wang and Hai Zhao

Shallow Discourse Parsing with Syntactic and (a Few) Semantic Features
Shubham Mukherjee, Abhishek Tiwari, Mohit Gupta and Anil Kumar Singh

The CLaC Discourse Parser at CoNLL-2015
Majid Laali, Elnaz Davoodi and Leila Kosseim

The DCU Discourse Parser for Connective, Argument Identification and Explicit Sense Classification
Longyue Wang, Chris Hokamp, Tsuyoshi Okita, Xiaojun Zhang and Qun Liu

The DCU Discourse Parser: A Sense Classification Task
Tsuyoshi Okita, Longyue Wang and Qun Liu

17:30-17:45 Session 8.b: Best Paper Award and Closing

Keynote Talk

On Spectral Graphical Models, and a New Look at Latent Variable Modeling in Natural Language Processing

Eric Xing

Carnegie Mellon University

epxing@cs.cmu.edu

Abstract

Latent variable and latent structure modeling, as widely seen in parsing systems, machine translation systems, topic models, and deep neural networks, represents a key paradigm in Natural Language Processing, where discovering and leveraging syntactic and semantic entities and relationships that are not explicitly annotated in the training set provide a crucial vehicle to obtain various desirable effects such as simplifying the solution space, incorporating domain knowledge, and extracting informative features. However, latent variable models are difficult to train and analyze in that, unlike fully observed models, they suffer from non-identifiability, non-convexity, and over-parameterization, which make them often hard to interpret, and tend to rely on local-search heuristics and heavy manual tuning.

In this talk, I propose to tackle these challenges using spectral graphical models (SGM), which view latent variable models through the lens of linear algebra and tensors. I show how SGMs exploit the connection between latent structure and low rank decomposition, and allow one to develop models and algorithms for a variety of latent variable problems, which unlike traditional techniques, enjoy provable guarantees on correctness and global optimality, can straightforwardly incorporate additional modern techniques such as kernels to achieve more advanced modeling power, and empirically offer a 1-2 orders of magnitude speed up over existing methods while giving comparable or better performance.

This is joint work with Ankur Parikh, Carnegie Mellon University.

Biography of the Speaker

Dr. Eric Xing is a Professor of Machine Learning in the School of Computer Science at Carnegie Mellon University, and Director of the CMU/UPMC Center for Machine Learning and Health. His principal research interests lie in the development of machine learning and statistical methodology, and large-scale computational system and architecture; especially for solving problems involving automated learning, reasoning, and decision-making in high-dimensional, multimodal, and dynamic possible worlds in artificial, biological, and social systems. Professor Xing received a Ph.D. in Molecular Biology from Rutgers University, and another Ph.D. in Computer Science from UC Berkeley. He servers (or served) as an associate editor of the Annals of Applied Statistics (AOAS), the Journal of American Statistical Association (JASA), the IEEE Transaction of Pattern Analysis and Machine Intelligence (PAMI), the PLoS Journal

of Computational Biology, and an Action Editor of the Machine Learning Journal (MLJ), the Journal of Machine Learning Research (JMLR). He was a member of the DARPA Information Science and Technology (ISAT) Advisory Group, a recipient of the NSF Career Award, the Sloan Fellowship, the United States Air Force Young Investigator Award, and the IBM Open Collaborative Research Award. He is the Program Chair of ICML 2014.

Keynote Talk

Does the Success of Deep Neural Network Language Processing Mean – Finally! – the End of Theoretical Linguistics?

Paul Smolensky

Johns Hopkins University

smolensky@jhu.edu

Abstract

Statistical methods in natural-language processing that rest on heavily empirically-based language learning specially those centrally deploying neural networks ave witnessed dramatic improvement in the past few years, and their success restores the urgency of understanding the relationship between (i) these neural/statistical language systems and (ii) the view of linguistic representation, processing, and structure developed over centuries within theoretical linguistics.

Two hypotheses concerning this relationship arise from our own mathematical and experimental results from past work, which we will present. These hypotheses can guide e will argue – important future research in the seemingly sizable gap separating computational linguistics from linguistic theories of human language acquisition. These hypotheses are:

1. The internal representational format used in deep neural networks for language – numerical vectors – is covertly an implementation of a system of discrete, symbolic, structured representations which are processed so as to optimally meet the demands of a symbolic grammar recognizable from the perspective of theoretical linguistics.
2. It will not be successes but rather the *failures* of future machine learning approaches to language acquisition which will be most telling for determining whether such approaches capture the crucial limitations on human language learning imitations, documented in recent artificial-grammar-learning experimental results, which support the nativist Chomskian hypothesis asserting that
    - reliably and efficiently learning human grammars from available evidence requires
    - that the hypothesis space entertained by the child concerning the set of possible (or likely) human languages
    - be limited by abstract, structure-based constraints;
    - these constraints can then also explain (in principle at least) the many robustly-respected universals observed in cross-linguistic typology.

This is joint work with Jennifer Culbertson, University of Edinburgh.

Biography of the Speaker

Paul Smolensky is the Krieger-Eisenhower Professor of Cognitive Science at Johns Hopkins University in Baltimore, Maryland, USA. He studies the mutual implications between the theories of neural computation and of universal grammar and has published in distinguished venues including Science and the Proceedings of the National Academy of Science USA. He received the David E. Rumelhart Prize for Outstanding Contributions to the Formal Analysis of Human Cognition (2005), the Chaire de Recherche Blaise Pascal (2008), and the Sapir Professorship of the Linguistic Society of America (2015). Primary results include:

- Contradicting widely-held convictions, (i) structured symbolic and (ii) neural network models of cognition are mutually compatible: formal descriptions of the same systems, the mind/brain, at (i) a highly abstract, and (ii) a more physical, level of description. His article n the proper treatment of connectionism1988) was until recently one of the 10-most cited articles in The Behavioral and Brain Sciences, itself the most-cited journal of all the behavioral sciences.
- That the theory of neural computation can in fact strengthen the theory of universal grammar is attested by the revolutionary impact in theoretical linguistics (within phonology in particular) of Optimality Theory, a neural- network-derived symbolic grammar formalism that he developed with Alan Prince (in a book widely released 1993, officially published 2004).
- The learnability theory for Optimality Theory was founded at nearly the same time as the theory itself, in joint work of Smolensky and his PhD student Bruce Tesar (TR 1993; article in the premier linguistic theory journal, Linguistic Inquiry 1998; MIT Press book 2000). This work laid the foundation upon which rests most of the flourishing formal theory of learning in Optimality Theory.
- There is considerable power in formalizing neural network computation as statistical inference/optimization within a dynamical system. Smolensky Harmony Theory (1981–6) analyzed network computation as Harmony Maximization (an independently-developed homologue to Hopfield nergy minimizationormulation) and first deployed principles of statistical inference for processing and learning in the bipartite network structure later to be known as the estricted Boltzmann Machinen the initial work on deep neural network learning (Hinton et al., 2006–).
- Powerful recursive symbolic computation can be achieved with massive parallelism in neural networks designed to process tensor product representations (TR 1987; journal article in Artificial Intelligence 1990). Related uses of the tensor product to structure numerical vectors is currently under rapid development in the field of distributional vector semantics.

Most recently, as argued in Part 1 of the talk, his work shows the value for theoretical and psycholinguistics of representations that share both the discrete structure of symbolic representations and the continuous variation of activity levels in neural network representations (initial results in an article in Cognitive Science by Smolensky, Goldrick & Mathis 2014).