ACL 2017

The 55th Annual Meeting of the Association for Computational Linguistics

Proceedings of the 2nd Workshop on Representation Learning for NLP

August 3, 2017
Vancouver, Canada

©2017 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-945626-62-3

Introduction

Welcome to the 2nd Workshop on Representation Learning for NLP (RepL4NLP), held on August 3, 2017 and hosted by the 55th Annual Meeting of the Association for Computational Linguistics (ACL) in Vancouver, Canada. The workshop is sponsored by DeepMind, Facebook AI Research, and Microsoft Research.

The 2nd Workshop on Representation Learning for NLP aims to continue the spirit of previously successful workshops at ACL/NAACL/EACL, namely RepL4NLP at ACL'16, VSM at NAACL'15 and CVSC at ACL'13/EACL'14/ACL'15, which focused on vector space models of meaning, compositionality, and the application of deep neural networks and spectral methods to NLP. It provides a forum for discussing recent advances on these topics, as well as future research directions in linguistically motivated vector-based models in NLP.

Organizers:

Phil Blunsom, DeepMind and Oxford University
Antoine Bordes, Facebook AI Research
Kyunghyun Cho, New York University
Shay Cohen, University of Edinburgh
Chris Dyer, DeepMind
Edward Grefenstette, DeepMind
Karl Moritz Hermann, DeepMind
Laura Rimell, University of Cambridge
Jason Weston, Facebook AI Research
Scott Yih, Microsoft Research

Program Committee:

Waleed Ammar
Eneko Agirre
Jacob Andreas
Isabelle Augenstein
Mohit Bansal
Marco Baroni
Jonathan Berant
Léon Bottou
Samuel Bowman
Junyoung Chung
Danqi Chen
Stephen Clark
Marco Damonte
Kevin Duh
Long Duong
Katrin Erk
Orhan Firat
Kevin Gimpel
Caglar Gulcehre
He He
Felix Hill
Mohit Iyyer
Kazuya Kawakami
Douwe Kiela
Jamie Ryan Kiros
Tomáš Kočiský
Angeliki Lazaridou
Omer Levy
Wang Ling
Shashi Narayan
Graham Neubig
Thien Huu Nguyen
Diarmuid Ó Séaghdha
Alexander Panchenko
Ankur Parikh
John Platt
Siva Reddy

Roi Reichart
Tim Rocktäschel
Hinrich Schütze
Holger Schwenk
Richard Socher
Mark Steedman
Karl Stratos
Sam Thomson
Yoshimasa Tsuruoka
Yulia Tsvetkov
Lyle Ungar
Oriol Vinyals
Andreas Vlachos
Dani Yogatama
Yi Yang

Table of Contents

Sense Contextualization in a Dependency-Based Compositional Distributional Model
Pablo Gamallo 1

Context encoders as a simple but powerful extension of word2vec
Franziska Horn 10

Machine Comprehension by Text-to-Text Neural Question Generation
Xingdi Yuan, Tong Wang, Caglar Gulcehre, Alessandro Sordoni, Philip Bachman, Saizheng Zhang, Sandeep Subramanian and Adam Trischler 15

Emergent Predication Structure in Hidden State Vectors of Neural Readers
Hai Wang, Takeshi Onishi, Kevin Gimpel and David McAllester 26

Towards Harnessing Memory Networks for Coreference Resolution
Joe Cheri and Pushpak Bhattacharyya 37

Combining Word-Level and Character-Level Representations for Relation Classification of Informal Text
Dongyun Liang, Weiran Xu and Yinge Zhao 43

Transfer Learning for Neural Semantic Parsing
Xing Fan, Emilio Monti, Lambert Mathias and Markus Dreyer 48

Modeling Large-Scale Structured Relationships with Shared Memory for Knowledge Base Completion
Yelong Shen, Po-Sen Huang, Ming-Wei Chang and Jianfeng Gao 57

Knowledge Base Completion: Baselines Strike Back
Rudolf Kadlec, Ondrej Bajgar and Jan Kleindienst 69

Sequential Attention: A Context-Aware Alignment Function for Machine Reading
Sebastian Brarda, Philip Yeres and Samuel Bowman 75

Semantic Vector Encoding and Similarity Search Using Fulltext Search Engines
Jan Rygl, Jan Pomikálek, Radim Řehůřek, Michal Růžička, Vít Novotný and Petr Sojka 81

Multi-task Domain Adaptation for Sequence Tagging
Nanyun Peng and Mark Dredze 91

Beyond Bilingual: Multi-sense Word Embeddings using Multilingual Context
Shyam Upadhyay, Kai-Wei Chang, Matt Taddy, Adam Kalai and James Zou 101

DocTag2Vec: An Embedding Based Multi-label Learning Approach for Document Tagging
Sheng Chen, Akshay Soni, Aasish Pappu and Yashar Mehdad 111

Binary Paragraph Vectors
Karol Grzegorczyk and Marcin Kurdziel 121

Representing Compositionality based on Multiple Timescales Gated Recurrent Neural Networks with Adaptive Temporal Hierarchy for Character-Level Language Models
Dennis Singh Moirangthem, Jegyung Son and Minho Lee 131

Learning Bilingual Projections of Embeddings for Vocabulary Expansion in Machine Translation
Pranava Swaroop Madhyastha and Cristina España-Bonet 139

Prediction of Frame-to-Frame Relations in the FrameNet Hierarchy with Frame Embeddings
Teresa Botschen, Hatem Mousselly Sergieh and Iryna Gurevych 146

Learning Joint Multilingual Sentence Representations with Neural Machine Translation
Holger Schwenk and Matthijs Douze 157

Transfer Learning for Speech Recognition on a Budget
Julius Kunze, Louis Kirsch, Ilia Kurenkov, Andreas Krug, Jens Johannsmeier and Sebastian Stober 168

Gradual Learning of Matrix-Space Models of Language for Sentiment Analysis
Shima Asaadi and Sebastian Rudolph 178

Improving Language Modeling using Densely Connected Recurrent Neural Networks
Fréderic Godin, Joni Dambre and Wesley De Neve 186

NewsQA: A Machine Comprehension Dataset
Adam Trischler, Tong Wang, Xingdi Yuan, Justin Harris, Alessandro Sordoni, Philip Bachman and Kaheer Suleman 191

Intrinsic and Extrinsic Evaluation of Spatiotemporal Text Representations in Twitter Streams
Lawrence Phillips, Kyle Shaffer, Dustin Arendt, Nathan Hodas and Svitlana Volkova 201

Rethinking Skip-thought: A Neighborhood based Approach
Shuai Tang, Hailin Jin, Chen Fang, Zhaowen Wang and Virginia de Sa 211

A Frame Tracking Model for Memory-Enhanced Dialogue Systems
Hannes Schulz, Jeremie Zumer, Layla El Asri and Shikhar Sharma 219

Plan, Attend, Generate: Character-Level Neural Machine Translation with Planning
Caglar Gulcehre, Francis Dutil, Adam Trischler and Yoshua Bengio 228

Does the Geometry of Word Embeddings Help Document Classification? A Case Study on Persistent Homology-Based Representations
Paul Michel, Abhilasha Ravichander and Shruti Rijhwani 235

Adversarial Generation of Natural Language
Sandeep Subramanian, Sai Rajeswar, Francis Dutil, Chris Pal and Aaron Courville 241

Deep Active Learning for Named Entity Recognition
Yanyao Shen, Hyokun Yun, Zachary Lipton, Yakov Kronrod and Animashree Anandkumar . . . 252

Learning when to skim and when to read
Alexander Johansen and Richard Socher 257

Learning to Embed Words in Context for Syntactic Tasks
Lifu Tu, Kevin Gimpel and Karen Livescu 265

Workshop Program

Thursday, August 3, 2017

09:30–09:45 Welcome and Opening Remarks

09:45–10:30 Keynote Session

09:45–10:30 *Learning Joint Embeddings of Vision and Language*
Sanja Fidler

A successful autonomous system needs to not only understand the visual world but also communicate its understanding with humans. To make this possible, language can serve as a natural link between high level semantic concepts and low level visual perception. In this talk, I'll discuss recent work in the domain of vision and language, covering topics such as image/video captioning and retrieval, and question-answering. I'll also talk about our recent work on task execution via language instructions.

10:30–11:00 Coffee Break

11:00–12:30 Keynote Session

11:00–11:45 *Learning Representations of Social Meaning*
Jacob Eisenstein

Language plays a critical role in structuring human relationships, while marking social properties of the speaker/writer, audience, and communicative situation. With the increasing availability of big social media datasets, computational linguists have begun to join with sociolinguists in working to elucidate language's social dimension. However, this promising synthesis is threatened by a theoretical mismatch between these two disciplines. Much of the research in the emerging field of computational sociolinguistics involves social-theoretical models that uncritically assign individuals to broad categories such as man/woman, black/white, northern/southern, and urban/rural. Meanwhile, sociolinguists have worked for decades to elaborate a more nuanced view of identity and social meaning, but it has proven difficult to reconcile these rich theoretical models with scalable quantitative research methods. In this talk, I will ask whether representation learning can help to bridge this gap. The key idea is to use learned representations to mediate between linguistic data and socially relevant metadata. I will describe applications of this basic approach in the context of clustering, latent variable models, and neural networks, with applications to gender, multi-community studies, and social network analysis.

Thursday, August 3, 2017 (continued)

11:45–12:30 *Representations in the Brain*
Alona Fyshe

What can the brain tell us about computationally-learned representations of words, phrases and beyond? And what can those computational representations tell us about the brain? In this talk I will describe several brain imaging experiments that explore the representation of language meaning in the brain, and relate those brain representations to computationally learned representations of language meaning.

12:30–14:00 Lunch

14:00–14:45 Keynote Session

14:00–14:45 *"A million ways to say I love you" or Learning to Paraphrase with Neural Machine Translation*
Mirella Lapata

Recognizing and generating paraphrases is an important component in many natural language processing applications. A well-established technique for automatically extracting paraphrases leverages bilingual corpora to find meaning-equivalent phrases in a single language by "pivoting" over a shared translation in another language. In the first part of the talk I will revisit bilingual pivoting in the context of neural machine translation and present a paraphrasing model based purely on neural networks. The proposed model represents paraphrases in a continuous space, estimates the degree of semantic relatedness between text segments of arbitrary length, and generates paraphrase candidates for any source input. In the second part of the talk I will illustrate how neural paraphrases can be seamlessly integrated in models of question answering and summarization, achieving competitive results across datasets and languages.

14:45–15:00 Best Paper Session

Thursday, August 3, 2017 (continued)

15:00–16:30 Poster Session, including Coffee Break

Sense Contextualization in a Dependency-Based Compositional Distributional Model
Pablo Gamallo

Context encoders as a simple but powerful extension of word2vec
Franziska Horn

Active Discriminative Text Representation Learning
Ye Zhang, Matthew Lease and Byron Wallace

Using millions of emoji occurrences to pretrain any-domain models for detecting emotion, sentiment and sarcasm
Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan and Sune Lehmann

Evaluating Layers of Representation in Neural Machine Translation on Syntactic and Semantic Tagging
Yonatan Belinkov, Lluís Màrquez, Hassan Sajjad, Nadir Durrani, Fahim Dalvi and James Glass

Machine Comprehension by Text-to-Text Neural Question Generation
Xingdi Yuan, Tong Wang, Caglar Gulcehre, Alessandro Sordoni, Philip Bachman, Saizheng Zhang, Sandeep Subramanian and Adam Trischler

Emergent Predication Structure in Hidden State Vectors of Neural Readers
Hai Wang, Takeshi Onishi, Kevin Gimpel and David McAllester

Towards Harnessing Memory Networks for Coreference Resolution
Joe Cheri and Pushpak Bhattacharyya

Combining Word-Level and Character-Level Representations for Relation Classification of Informal Text
Dongyun Liang, Weiran Xu and Yinge Zhao

Regularized Topic Models for Sparse Interpretable Word Embeddings
Anna Potapenko and Artem Popov

Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings
Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama and Adam T. Kalai

Transfer Learning for Neural Semantic Parsing
Xing Fan, Emilio Monti, Lambert Mathias and Markus Dreyer

MUSE: Modularizing Unsupervised Sense Embeddings
Guang-He Lee and Yun-Nung Chen

Modeling Large-Scale Structured Relationships with Shared Memory for Knowledge Base Completion
Yelong Shen, Po-Sen Huang, Ming-Wei Chang and Jianfeng Gao

Knowledge Base Completion: Baselines Strike Back
Rudolf Kadlec, Ondrej Bajgar and Jan Kleindienst

Sequential Attention: A Context-Aware Alignment Function for Machine Reading
Sebastian Brarda, Philip Yeres and Samuel Bowman

Semantic Vector Encoding and Similarity Search Using Fulltext Search Engines
Jan Rygl, Jan Pomikálek, Radim Řehůřek, Michal Růžička, Vít Novotný and Petr Sojka

Multi-task Domain Adaptation for Sequence Tagging
Nanyun Peng and Mark Dredze

Beyond Bilingual: Multi-sense Word Embeddings using Multilingual Context
Shyam Upadhyay, Kai-Wei Chang, Matt Taddy, Adam Kalai and James Zou

DocTag2Vec: An Embedding Based Multi-label Learning Approach for Document Tagging
Sheng Chen, Akshay Soni, Aasish Pappu and Yashar Mehdad

Binary Paragraph Vectors
Karol Grzegorczyk and Marcin Kurdziel

Representing Compositionality based on Multiple Timescales Gated Recurrent Neural Networks with Adaptive Temporal Hierarchy for Character-Level Language Models
Dennis Singh Moirangthem, Jegyung Son and Minho Lee

Learning Bilingual Projections of Embeddings for Vocabulary Expansion in Machine Translation
Pranava Swaroop Madhyastha and Cristina España-Bonet

Learning to Compose Words into Sentences with Reinforcement Learning
Dani Yogatama, Phil Blunsom, Chris Dyer, Edward Grefenstette and Wang Ling

Prediction of Frame-to-Frame Relations in the FrameNet Hierarchy with Frame Embeddings
Teresa Botschen, Hatem Mousselly Sergieh and Iryna Gurevych

Learning Joint Multilingual Sentence Representations with Neural Machine Translation
Holger Schwenk and Matthijs Douze

Transfer Learning for Speech Recognition on a Budget
Julius Kunze, Louis Kirsch, Ilia Kurenkov, Andreas Krug, Jens Johannsmeier and Sebastian Stober

Gradual Learning of Matrix-Space Models of Language for Sentiment Analysis
Shima Asaadi and Sebastian Rudolph

Improving Language Modeling using Densely Connected Recurrent Neural Networks
Fréderic Godin, Joni Dambre and Wesley De Neve

NewsQA: A Machine Comprehension Dataset
Adam Trischler, Tong Wang, Xingdi Yuan, Justin Harris, Alessandro Sordoni, Philip Bachman and Kaheer Suleman

Intrinsic and Extrinsic Evaluation of Spatiotemporal Text Representations in Twitter Streams
Lawrence Phillips, Kyle Shaffer, Dustin Arendt, Nathan Hodas and Svitlana Volkova

Rethinking Skip-thought: A Neighborhood based Approach
Shuai Tang, Hailin Jin, Chen Fang, Zhaowen Wang and Virginia de Sa

A Frame Tracking Model for Memory-Enhanced Dialogue Systems
Hannes Schulz, Jeremie Zumer, Layla El Asri and Shikhar Sharma

Plan, Attend, Generate: Character-Level Neural Machine Translation with Planning
Caglar Gulcehre, Francis Dutil, Adam Trischler and Yoshua Bengio

Does the Geometry of Word Embeddings Help Document Classification? A Case Study on Persistent Homology-Based Representations
Paul Michel, Abhilasha Ravichander and Shruti Rijhwani

Adversarial Generation of Natural Language
Sandeep Subramanian, Sai Rajeswar, Francis Dutil, Chris Pal and Aaron Courville

Deep Active Learning for Named Entity Recognition
Yanyao Shen, Hyokun Yun, Zachary Lipton, Yakov Kronrod and Animashree Anandkumar

The Coadaptation Problem when Learning How and What to Compose
Andrew Drozdov and Samuel Bowman

Learning when to skim and when to read
Alexander Johansen and Richard Socher

Learning to Embed Words in Context for Syntactic Tasks
Lifu Tu, Kevin Gimpel and Karen Livescu

16:30–17:30 Panel Discussion

17:30–17:40 Closing Remarks