Introduction to the Special Section on Continuous Space and Related Methods in Natural Language Processing

Natural Language Processing (NLP) aims to analyze, understand, and generate languages that humans use naturally. Significant progress in NLP has been achieved in recent years, addressing important and practical real-world problems, and enabling mass deployment of large-scale systems. New machine learning paradigms such as deep learning and continuous space methods have contributed to inferring language patterns from increasingly large real-world data and to making predictions about new data more accurate.

One of the challenges in NLP is to represent language in a form that can be processed effectively by computing algorithms. Words in sequence are traditionally treated as discrete symbols, which has advantages and limitations. Research on continuous space methods provides a promising alternative by describing words and their semantic and syntactic relationships in a different way. In continuous space language modeling, we represent words with real-valued vectors. In this way, conditional probability distributions of words can be learned and expressed as smooth functions of these vectors; similar words are therefore described as neighbors in the continuous space. A Neural Network Language Model is a typical example of such continuous space methods.

Building on the success of acoustic and statistical language modeling, research on artificial (deep) neural networks, and continuous space models in general, has seen significant progress in mitigating data sparseness, incorporating longer contexts, and modeling morphological, syntactic and semantic relationships across words. As a result, continuous space models are now embedded in many state-of-the-art speech recognition, machine translation and spoken language understanding systems.

This special section includes ten papers that discuss some latest findings on research problems related to the application of continuous space and related models in NLP. They were selected by five Guest Editors through a regular review process.

Recent studies show that neural networks can effectively learn vector representations for words. These vector representations precisely capture the syntactic and semantic information of words. In [1] "Syntactic and Semantic Features for Code-Switching Factored Language Models," Heike Adel, Ngoc Thang Vu, Katrin Kirchhoff, Dominic Telaar, and Tanja Schultz study using a recurrent neural network learned word embedding to improve the quality of factorized language models for code-switching speech. In [2] "Graph-based Lexicon Regularization for PCFG with Latent Annotations," Xiaodong Zen, Derek F. Wong, Lidia S. Chao, and Isabel Trancoso, study a graph propagation technique for probabilistic context-free grammar with latent annotations, where a neural word representation approach was adopted to capture syntactic and semantic information from word co-occurrence statistics. In [3] "Distributed Feature Representations for Dependency Parsing," Wenliang Chen, Yue Zhang and Min Zhang study how the features of dependency parsing models can be represented as distributed vectors in a continuous space, and then be utilized in a graph-based dependency parsing model. The experimental results on benchmarking data show that the parsers with the distributed feature representations significantly outperform several very strong baseline parsers with different settings. In [4] "Learning Semantic Hierarchies: A Continuous Vector Space Approach," Ruiji Fu, Jiang Guo, Bing Qin, Wanxiang Che, Haifeng Wang, and Ting Liu propose a novel and effective method for the construction of semantic hierarchies based on continuous vector representation of words, which can be used to measure the semantic relationship between words. The authors identify whether a candidate word pair has hypernym–hyponym relation by using the word-embedding-based semantic projections between words and their hypernyms.

Continuous space methods also offer new solutions to traditional research problems such as Machine Translation. In [5], "Adequacy–Fluency Metrics: Evaluating MT in the Continuous Space Model Framework," Rafael E. Banchs, Luis F. D'Haro, and Haizhou Li study how continuous space representations can be used for assessing machine translation quality even without the need for reference translations. This demonstrates the still hidden potential of continuous space representations for modeling meaning and semantics across the cognitive frontiers of language itself. In [6], "Topic-Based Coherence Modeling for Statistical Machine Translation," Deyi Xiong, Xing Wang and Min Zhang aim at capturing the structure of documents with the objective of improving the quality of statistical machine translation. This work adopts the linguistic notion of coherence as a "continuity of senses." Hence, coherence is expressed probabilistically as a continuous sense transition over sentences within a document. The presented approach associates a topic to each sentence in a coherent source document, and the topic chain is projected onto the target document. The topic-based coherence model brought significant improvements in statistical machine translation.

Neural network approaches have proven successful in addressing practical issues in NLP. Language modeling for low

Digital Object Identifier 10.1109/TASLP.2015.2405131

resource languages is an important research problem. Addressing the problem, Brian Hutchinson, Mari Ostendorf, and Maryam Fazel propose a new exponential language model in [7] "A Sparse Plus Low Rank Exponential Language Model." The model learns the sparse and low-rank structure in the language, while the low rank matrix corresponds to a continuous-space vector representation of words and word histories. Experiments show that this model leads to superior performance in scenarios of low resource languages and domain adaptation. In [8] "Deep Learning Framework with Confused Sub-set Resolution Architecture for Automatic Arabic Diacritization," Mohsen A. A. Rashwan, Ahmad A. Al Sallab, Hazem M. Raafat, and Ahmed Rafea study a deep learning framework for automatic Arabic diacritization. The Arabic diacritics restoration task is formulated as a pattern classification problem where a deep network architecture serves as the classifier with diacritics as the output and the Modern Standard Arabic transcripts as the input. The experiments demonstrate the clear advantage of deep learning framework over other approaches.

A systematic study of different neural network architectures helps us understand how neural network methods are superior in estimating probability distributions over word sequences. In [9] "From Feedforward to LSTM Neural Network Language Models," Martin Sundermeyer, Hermann Ney, and Ralf Schluter, compare conventional count based language models against different neural network architectures, including feedforward, recurrent and long short term memory networks. Their comparison, which is carried out on two large vocabulary tasks, points out the benefits of neural network on automatic speech recognition performance, but also many efficiency issues concerning the integration of neural networks into search algorithms. In [10] "Using Recurrent Neural Network for Slot Filling in Spoken Language Understanding," Grégoire Mesnil, Yann Dauphin, Kaisheng Yao, Yoshua Bengio, Li Deng, Dilek Hakkani-Tur, Xiaodong He, Larry Heck, Gokhan Tur, Dong Yu, and Geoffrey Zweig explore the use of recurrent neural networks (RNNs) for the task of slot filling in spoken language understanding. The authors explore multiple RNN architectures that can efficiently deal with temporal dependencies, and evaluate these models on the public ATIS air-travel benchmark, as well as two other understanding tasks dealing with entertainment and movies. The RNN slot filling models are found to be highly competitive with a conditional random field (CRF) baseline.

In response to the increased interest in continuous space and related methods, IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (T-ASLP) published a special section in the January issue of Volume 20, in 2012, on Deep Learning for Speech and Language Processing, with five papers covering topics in speech processing. This special section furthers the effort by featuring ten papers all in natural language processing, which share the enthusiasm with the text processing community.

Finally, we hope this special section will further stimulate interest in continuous space and related methods in natural language processing. We would like to thank all the authors who submitted their papers, and the reviewers for providing valuable feedbacks that have greatly improved the articles. Special thanks go to Li Deng, the former Editor-in-Chief, as well as to Mari Ostendorf, the former VP of IEEE SPS Publication Board for their guidance. We are also grateful to Bowen Chen and Jim Glass, the Associate Editors of T-ASLP for their assistance in the editorial process.

Haizhou Li, *Lead Guest Editor*
Institute for Infocomm Research
Singapore, 138632

Marcello Federico, *Guest Editor*
Fondazione Bruno Kessler
38122 Trento, Italy

Xiaodong He, *Guest Editor*
Microsoft Research
Redmond , WA 98052 USA

Helen Meng, *Guest Editor*
The Chinese University of Hong Kong
Hong Kong, China

Isabel Trancoso, *Guest Editor*
INESC-ID and Instituto Superior Técnico
University of Lisbon
Lisbon, 1649-004 Portugal

REFERENCES

[1] H. Adel, N. Thang Vu, K. Kirchhoff, D. Telaar, and T. Schultz, "Syntactic and semantic features for code-switching factored language models," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 431–440, Mar. 2015.

[2] X. Zen, D. F. Wong, L. S. Chao, and I. Trancoso, "Graph-based lexicon regularization for PCFG with latent annotations," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 441–450, Mar. 2015.

[3] W. Chen, Y. Zhang, and M. Zhang, "Distributed feature representations for dependency parsing," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 451–460, Mar. 2015.

[4] R. Fu, J. Guo, B. Qin, W. Che, H. Wang, and T. Liu, "Learning semantic hierarchies: A continuous vector space approach," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 461–471, Mar. 2015.

[5] R. E. Banchs, L. F. D'Haro, and L. Haizhou, "Adequacy–fluency metrics: Evaluating MT in the continuous space model framework," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 472–482, Mar. 2015.

[6] D. Xiong, X. Wang, and M. Zhang, "Topic-based coherence modeling for statistical machine translation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 483–493, Mar. 2015.

[7] B. Hutchinson, M. Ostendorf, and M. Fazel, "A sparse plus low rank exponential language model," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 494–504, Mar. 2015.

[8] M. A. A. Rashwan, A. A. Al Sallab, H. M. Raafat, and A. Rafea, "Deep learning framework with confused sub-set resolution architecture for automatic Arabic diacritization," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 505–516, Mar. 2015.

[9] M. Sundermeyer, H. Ney, and R. Schluter, "From feedforward to LSTM neural network language models," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 517–529, Mar. 2015.

[10] G. Mesnil, Y. Dauphin, K. Yao, Y. Bengio, L. Deng, D. Hakkani-Tur, X. He, L. Heck, G. Tur, D. Yu, and G. Zweig, "Using recurrent neural network for slot filling in spoken language understanding," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 3, pp. 530–539, Mar. 2015.

Haizhou Li (M'91–SM'01–F'14) received the B.Sc., M.Sc., and Ph.D. degree in electrical and electronic engineering from South China University of Technology, Guangzhou, China, in 1984, 1987, and 1990, respectively. Dr. Li is currently the Principal Scientist, Department Head of Human Language Technology in the Institute for Infocomm Research (I^2R), Singapore. He is also an Adjunct Professor at the National University of Singapore and a Conjoint Professor at the University of New South Wales, Australia. His research interests include automatic speech recognition, speaker and language recognition, and natural language processing.

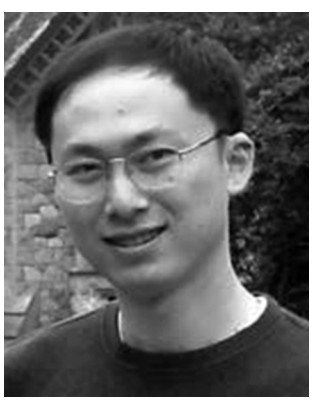
Prior to joining I^2R, he taught in the University of Hong Kong (1988–1990) and South China University of Technology (1990–1994). He was a Visiting Professor at CRIN in France (1994–1995), a Research Manager at the Apple-ISS Research Centre (1996–1998), a Research Director in Lernout & Hauspie Asia Pacific (1999–2001), and the Vice President in InfoTalk Corp. Ltd. (2001–2003).

Dr. Li is currently the Editor-in-Chief of IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (2015–2017). He has served in the Editorial Board of *Computer Speech and Language* (2012–2014). He is an elected Member of IEEE Speech and Language Processing Technical Committee (2013–2015), the Vice President of the International Speech Communication Association (2013–2014), President of Asia Pacific Signal and Information Processing Association (2015–2016). He was the General Chair of ACL 2012 and INTERSPEECH 2014.

Dr. Li is a Fellow of IEEE. He was a recipient of the National Infocomm Award 2002 and the President's Technology Award 2013 in Singapore. He was named one the two Nokia Visiting Professors in 2009 by the Nokia Foundation.

Marcello Federico heads the HLT-MT research unit at Fondazione Bruno Kessler and is Adjunct Professor at the ICT Doctorate School of the University of Trento and at the ISIT Institute in Trento His research expertise is in statistical machine translation, spoken language translation, statistical language modeling, information retrieval, and speech recognition. In these areas, he has co-authored over 150 scientific papers, contributed in 20 international and national projects, and co-developed widely used software packages for machine translation and language modeling.

He has served in the program committees of all major international conferences in the field of human language technology. He has been member of the board of the European Association for Machine Translation [2010-2014], regional officer of ACL's Special Interest Group on Machine Translation [2010–2014]. He has been Editor-in-Chief of the *ACM Transactions on Speech and Language Processing* [2007–2013], Associate Editor for ACM TSLP [2005–2007] and Foundations and Trends in Information Retrieval [2005–2009]. Since 2014 he has been Senior Editor for the IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING.



Xiaodong He (M'03–SM'08) is a Researcher with Microsoft Research, Redmond, WA, USA. He is also an Affiliate Professor in electrical engineering at the University of Washington, Seattle, WA, USA. His research interests include deep learning, information retrieval, natural language understanding, machine translation, and speech recognition. Dr. He and his colleagues have developed entries that obtained No. 1 place in the 2008 NIST Machine Translation Evaluation (NIST MT) and the 2011 International Workshop on Spoken Language Translation Evaluation (IWSLT), both in Chinese–English translation, respectively. He serves as Associate Editor of the *IEEE Signal Processing Magazine* and IEEE SIGNAL PROCESSING LETTERS, as Guest Editors of IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING for the Special Issue on Continuous-space and related methods in natural language processing, and Area Chair of NAACL2015. He also served as GE for several IEEE Journals, and served in organizing committees and program committees of major speech and language processing conferences in the past. He is a senior member of IEEE and a member of ACL.

Helen M. Meng (M'98–SM'09–F'13) received the S. B., S. M., and Ph.D. degrees, all in electrical engineering, from the Massachusetts Institute of Technology, Cambridge. She has been a Research Scientist at the MIT Spoken Language Systems Group, where she worked on multilingual conversational systems. She joined The Chinese University of Hong Kong in 1998, where she is currently Professor and Chairman of the Department of Systems Engineering and Engineering Management. She was also the past Associate Dean of Research of the Faculty of Engineering. In 1999, she established the Human–Computer Communications Laboratory at CUHK and serves as director. In 2005, she established the Microsoft-CUHK Joint Laboratory for Human-Centric Computing and Interface Technologies, which was upgraded to MoE-Microsoft Key Laboratory in 2008, and serves as Co-Director. She is also Co-Director of the Tsinghua-CUHK Joint Research Center for Media Sciences, Technologies and Systems. Her research interest is in the area of human–computer interaction via multimodal and multilingual spoken language systems, as well as translingual speech retrieval technologies. She was the Editor-in-Chief of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (2009–2011). She is a Fellow of IEEE.

Isabel Trancoso (SM'03–F'11)received the Licenciado, Mestre, Doutor, and Agregado degrees in electrical and computer engineering from Instituto Superior Técnico, University of Lisbon, Portugal, in 1979, 1984, 1987, and 2002, respectively. She has been a Lecturer at this University since 1979, having coordinated the EEC course for 6 years. She is currently a Full Professor, Chair of the EEC Department where she teaches speech processing courses. She is also a Senior Researcher at INESCID Lisbon, having launched the speech processing group, now restructured as L2F, in 1990. Her first research topic was medium-to-low bit rate speech coding. In October 1984–June 1985, she worked on this topic at AT&T Bell Laboratories, Murray Hill, New Jersey. Her current scope is much broader, encompassing many areas in spoken language processing, with a special emphasis on the Portuguese language. She was a member of the ISCA (International Speech Communication Association) Board (1993–1998), the IEEE Speech Technical Committee and the Permanent Council for the Organization of the International Conferences on Spoken Language Processing. She was elected Editor in Chief of the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING (2003–2005), Member-at-Large of the IEEE Signal Processing Society Board of Governors (2006–2008), Vice-President of ISCA (2005–2007) and President of ISCA (2007–2011). She chaired the Organizing Committee of the INTERSPEECH'05 Conference that took place in September 2005, in Lisbon. She launched and coordinates the Dual Degree Ph.D. Program of the Carnegie Mellon Portugal Partnership in Language Technologies. She chaired the IEEE James L. Flanagan Speech and Audio Processing Award Committee, and is a member of the IEEE Fellows Committee, of the IEEE Publication Services and Products Board, and of the Editorial Board of the *IEEE Signal Processing Magazine*. She chairs the Distinguished Lecturer Selection Committee of ISCA and is part of the ISCA Advisory Committee. She is Vice-President of the European Language Resources Association. She received the IEEE Signal Processing Society Meritorious Service Award in 2009, was elevated to IEEE Fellow in 2011, and to ISCA Fellow in 2014.