TextGraphs-9

Graph-Based Methods for Natural Language Processing

Proceedings of the Workshop

October 29, 2014
Doha, Qatar

Production and Manufacturing by
Taberg Media Group AB
Box 94, 562 02 Taberg
Sweden

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 978-1-937284-96-1

Introduction to TextGraphs-9

Welcome to TextGraphs, the workshop on Graph-based Methods for Natural Language Processing. The ninth edition of the workshop is being organized on October 29, 2014, in conjunction with the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014 in Doha, Qatar.

For the past eight years, the series of TextGraphs workshops have exposed and encouraged the synergy between the field of Graph Theory and Natural Language Processing (NLP). The mix between the two started small, with graph theoretical framework providing efficient and elegant solutions for NLP applications that focused on single documents for part-of-speech tagging, word sense disambiguation and semantic role labelling, got progressively larger with ontology learning and information extraction from large text collections, and have reached web scale through the new fields of research that focus on information propagation in social networks, rumor proliferation, e-reputation, multiple entity detection, language dynamics learning and future events prediction to name but a few.

The ninth edition of the TextGraphs workshop would be a new step in the series, focused on issues and solutions for large-scale graphs, such as those derived for web-scale knowledge acquisition or social networks. We encourage the description of novel NLP problems or applications that have emerged in recent years which can be addressed with graph-based solutions, as well as novel graph-based methods that can be applied to known NLP tasks. Continuing to bring together researchers interested in Graph Theory applied to Natural Language Processing provides an environment for further integration of graph-based solutions into NLP tasks. A deeper understanding of new theories of graph-based algorithms is likely to help create new approaches and widen the usage of graphs for NLP applications.

This volume contains papers accepted for presentation at the workshop. We issued calls for regular papers, short papers, position papers, and demos. After careful review by the program committee, 6 regular papers and 2 short papers were accepted for presentation. The accepted papers address varied problems – from theoretical and general considerations, to NLP and real-world applications - through interesting variations to known and also novel graph-based methods.

We are lucky to have two excellent invited speakers for this year's event. We thank Prof. Mohammed J. Zaki and Partha Talukdar for their enthusiastic acceptance to our invitation.

Finally, we are thankful to the members of the program committee for their valuable and high quality reviews. All submissions have benefited from their expert feedback. Their timely contribution was the basis for accepting an excellent list of papers and making this edition of TextGraphs a success.

V.G.Vinod Vydiswaran, Amarnag Subramanya, Gabor Melli, and Irina Matveeva

October 2014

Workshop Organizers:

V.G.Vinod Vydiswaran, University of Michigan (USA)
Amarnag Subramanya, Google (USA)
Gabor Melli, VigLink (USA)
Irina Matveeva, NexLP (USA)

Program Committee:

Asif Ekbar, Indian Institute of Technology, Patna (India)
Filip Ginter, University of Turku (Finland)
Rada Mihalcea, University of Michigan (USA)
Animesh Mukherjee, Indian Institute of Technology, Kharagpur (India)
Philippe Muller, Paul Sabatier University (France)
Preslav Nakov, Qatar Computing Research Institute (Qatar)
Günter Neumann, DFKI, Saarbrücken (Germany)
Arzucan Özgür, Bogazici University (Turkey)
Simone Paolo Ponzetto, University of Mannheim (Germany)
Delip Rao, Twitter (USA)
Martin Riedl, Darmstadt University of Technology (Germany)
Fabio Massimo Zanzotto, University of Rome (Italy)

Invited Speakers:

Mohammed J. Zaki, Rensselaer Polytechnic Institute (USA)
Partha Talukdar, Indian Institute of Science (India)

Table of Contents

Normalized Entity Graph for Computing Local Coherence
Mohsen Mesgar and Michael Strube 1

Exploiting Timegraphs in Temporal Relation Classification
Natsuda Laokulrat, Makoto Miwa and Yoshimasa Tsuruoka 6

Multi-document Summarization Using Bipartite Graphs
Daraksha Parveen and Michael Strube 15

A Novel Two-stage Framework for Extracting Opinionated Sentences from News Articles
Pujari Rajkumar, Swara Desai, Niloy Ganguly and Pawan Goyal 25

Constructing Coherent Event Hierarchies from News Stories
Goran Glavaš and Jan Šnajder 34

Semi-supervised Graph-based Genre Classification for Web Pages
Noushin Rezapour Asheghi, Katja Markert and Serge Sharoff 39

The Modular Community Structure of Linguistic Predication Networks
Aaron Gerow and James Evans 48

From Visualisation to Hypothesis Construction for Second Language Acquisition
Shervin Malmasi and Mark Dras 56

Conference Program

Wednesday, October 29, 2014

Session 1

09:00–09:10 *Welcome and Introduction*
The organizers

09:10–10:15 *Keynote Talk*
Prof. Mohammed J. Zaki

10:15–10:30 *Normalized Entity Graph for Computing Local Coherence*
Mohsen Mesgar and Michael Strube

10:30–11:00 ***Coffee break***

Session 2

11:00–11:25 *Exploiting Timegraphs in Temporal Relation Classification*
Natsuda Laokulrat, Makoto Miwa and Yoshimasa Tsuruoka

11:25–11:50 *Multi-document Summarization Using Bipartite Graphs*
Daraksha Parveen and Michael Strube

11:50–12:15 *A Novel Two-stage Framework for Extracting Opinionated Sentences from News Articles*
Pujari Rajkumar, Swara Desai, Niloy Ganguly and Pawan Goyal

12:15–12:30 *Constructing Coherent Event Hierarchies from News Stories*
Goran Glavaš and Jan Šnajder

12:30–14:00 ***Lunch***

Wednesday, October 29, 2014 (continued)

Session 3

14:00–15:05 *Invited Talk*
Prof. Partha Talukdar

15:05–15:30 *Semi-supervised Graph-based Genre Classification for Web Pages*
Noushin Rezapour Asheghi, Katja Markert and Serge Sharoff

15:30–16:00 ***Coffee break***

Session 4

16:00–16:25 *The Modular Community Structure of Linguistic Predication Networks*
Aaron Gerow and James Evans

16:25–16:50 *From Visualisation to Hypothesis Construction for Second Language Acquisition*
Shervin Malmasi and Mark Dras

16:50–17:00 *Conclusion*
The organizers