NAACL HLT 2009

Integer Linear Programming for Natural Language Processing

Proceedings of the Workshop

June 4, 2009
Boulder, Colorado

Production and Manufacturing by
Omnipress Inc.
2600 Anderson Street
Madison, WI 53707
USA

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 978-1-932432-35-0

Introduction

We are pleased to present the proceedings of the Workshop on Integer Linear Programming for Natural Language Processing, held at NAACL HLT 2009 in Boulder, Colorado.

Integer Linear Programming (ILP) has recently attracted much attention within the NLP community. Formulating problems using ILP has several advantages. It allows us to focus on the modelling of problems, rather than engineering new search algorithms; provides the opportunity to incorporate generic global constraints; and guarantees exact inference. This and the availability of off-the-shelf solvers has led to a large variety of natural language processing tasks being formulated in the ILP framework, including semantic role labelling, syntactic parsing, summarisation and joint information extraction.

The use of ILP brings many benefits and opportunities but there are still challenges for the community; these include: formulations of new applications, dealing with large-scale problems and understanding the interaction between learning and inference at training and decision time. The purpose of this workshop was to bring together researchers interested in exploiting ILP for NLP applications and tackling the issues involved. We solicited full length papers, short papers and two-page abstracts with the purpose of providing a discussion on many topics.

We are grateful to the program committee for providing thoughtful and helpful reviews of the submitted papers. We also thank our invited speakers, Dan Roth, Andre Martins and Noah Smith for presenting their noteworthy work to the community.

We hope that you enjoy the workshop and these proceedings.

James Clarke
Sebastian Riedel

Organizers:

James Clarke, University of Illinois at Urbana-Champaign
Sebastian Riedel, University of Tokyo/DBCLS

Program Committee:

Ming-Wei Chang, University of Illinois at Urbana-Champaign
Hal Daume III, University of Utah
Pascal Denis, INRIA Paris-Rocquencourt
Mark Dras, Macquarie University
Katja Filippova, EML Research gGmbH
Jenny Rose Finkel, Stanford University
Manfred Klenner, University of Zurich
Kevin Knight, University of Southern California
Mirella Lapata, University of Edinburgh
Daniel Marcu, University of Southern California
Ryan McDonald, Google Research
Ivan Meza-Ruiz, University of Edinburgh
Nick Rizzolo, University of Illinois at Urbana-Champaign
Dan Roth, University of Illinois at Urbana-Champaign
Hiroya Takamura, Tokyo Institute of Technology
Scott Yih, Microsoft Research

Invited Speakers:

Andre Martins, Carnegie Mellon University
Dan Roth, University of Illinois at Urbana-Champaign
Noah Smith, Carnegie Mellon University

Table of Contents

Conference Program

Thursday, June 4, 2009

9:20–9:30 Opening Remarks

9:30–10:30 Invited Talk by Dan Roth

10:30–11:00 Coffee break

11:00–11:30 *Summarization with a Joint Model for Sentence Extraction and Compression*
Andre Martins and Noah A. Smith

11:30–12:00 *A Scalable Global Model for Summarization*
Dan Gillick and Benoit Favre

12:00–12:30 *Bounding and Comparing Methods for Correlation Clustering Beyond ILP*
Micha Elsner and Warren Schudy

12:30–2:00 Lunch break

2:00–2:45 Invited Talk by Andre Martins and Noah Smith

2:45–3:15 *A New Objective Function for Word Alignment*
Tugba Bodrumlu, Kevin Knight and Sujith Ravi

3:15–3:30 *A Constraint Programming Approach to Probabilistic Syntactic Processing*
Irene Langkilde-Geary

3:30–4:00 Coffee break

4:00–5:00 Discussion