

SPONSOR:

CLIF (Computational Linguistics in Flanders)

INVITED SPEAKER:

Steven Abney, University of Michigan

PROGRAM COMMITTEE:

Walter Daelemans, University of Antwerp and Tilburg University, co-chair
Miles Osborne, University of Edinburgh, co-chair

Thorsten Brants, Google Inc.
Claire Cardie, Cornell University
James Cussens, University of York
Ido Dagan, Bar Ilan University
Diane Litman, University of Pittsburgh
Yuji Matsumoto, Nara Institute of Science and Technology
Rada Mihalcea, University of North Texas
Raymond Mooney, University of Texas at Austin
John Nerbonne, Groningen University
Hwee-Tou Ng, National University of Singapore
Grace Ngai, The Hong Kong Polytechnic University
David Powers, Flinders University
Adwait Ratnaparkhi, Microsoft
Ellen Riloff, University of Utah
Dan Roth, University of Illinois at Urbana-Champaign
Erik Tjong Kim Sang, University of Antwerp
Antal van den Bosch, Tilburg University

Additional Reviewers:

Wen-Tau Yih, University of Illinois at Urbana-Champaign
Vivi Năstase, University of Ottawa
Yuk Wah Wong, University of Texas at Austin
Ruifang Ge, University of Texas at Austin
Martin Reynaert, University of Tilburg
Steve Clark, University of Edinburgh
Răzvan Bunescu, University of Texas at Austin
Véronique Hoste, University of Antwerp
Bart Decadt, University of Antwerp
Anja Höthker, University of Antwerp
Fien De Meulder, University of Antwerp

SHARED TASK COORDINATORS:

Erik Tjong Kim Sang, University of Antwerp
Fien De Meulder, University of Antwerp

CONFERENCE WEBSITE:

<http://cnts.uia.ac.be/conll2003/>

PREFACE

This volume contains the papers accepted for presentation at the 2003 conference on Computational Natural Language Learning (CoNLL-2003), a SIGNLL meeting held in Edmonton, Canada, May 31 and June 1 in conjunction with the HLT-NAACL Conference.

CoNLL-2003 is the seventh in a series of meetings organized by SIGNLL¹, the ACL special interest group on natural language learning. The CoNLL meeting is intended to address all aspects of computational natural language learning.

This year, we have a special theme on semi-supervised learning; the development of techniques for making use of the massively available unannotated text data in order to increase accuracy of trained language processing systems. Our keynote speaker, Steven Abney, fits this theme with a presentation about the Yarowsky algorithm.

As in the previous three editions of CoNLL, we have a shared task session, coordinated by Erik Tjong Kim Sang and Fien De Meulder. This year the task is named-entity recognition. For this session, training and test data were made available, and researchers were invited to apply their learning system to this task. Seventeen systems have participated. You will find descriptions of these systems and their results in these proceedings.

Of the 35 full papers submitted, the programme committee selected 17 papers, representative of the state of the art in computational language learning today. We are very grateful to our program committee and additional reviewers for the effort they put in reviewing the papers. We are also grateful to the HLT-NAACL-2003 conference organizers on whom we could rely for the local organization. Finally, we most gratefully acknowledge the support of our sponsor, the Computational Linguistics in Flanders research community (CLIF) for sponsoring the invited speaker.

Walter Daelemans & Miles Osborne (editors)
April 2003

¹<http://www.aclweb.org/signll/>