ACL 2007

PRAGUE

IWPT 2007

Proceedings of the 10th International Conference on Parsing Technologies

June 23–34, 2007
Prague, Czech Republic

The Association for Computational Linguistics

Production and Manufacturing by
Omnipress
2600 Anderson Street
Madison, WI 53704
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

Preface

Welcome to the Tenth International Conference on Parsing Technologies, IWPT 2007, in the beautiful city of Prague.

IWPT'07 continues the tradition of biennial workshops on parsing technology organized by SIGPARSE, the Special Interest Group on Parsing of the Association for Computational Linguistics (ACL). The first workshop, in Pittsburgh and Hidden Valley, was followed by workshops in Cancun (Mexico) in 1991; Tilburg (Netherlands) and Durbuy (Belgium) in 1993; Prague and Karlovy Vary (Czech Republic) in 1995; Boston/Cambridge (Massachusetts) in 1997; Trento (Italy) in 2000; Beijing (China) in 2001; Nancy (France) in 2003; and Vancouver (Canada) in 2005.

Over the years the IWPT Workshops have become the major forum for researchers in natural language parsing. They have also given rise to four books on parsing technologies.

For the first time this year, IWPT is organised as a co-located event with the main ACL conference and with EMNLP and many other workshops. We would like to thank Alon Lavie, Priscilla Rasmussen and the ACL committee and local organisers for their help and support in organising this event.

Parsing technologies are relevant for almost all applications in Natural Language Processing. We are fortunate to have Stuart Shieber from Harvard University as our invited speaker to explore the links between sychronous grammars and issues related to machine translation and parsing.

This year's programme features for the first time invited presentations by organisers of co-located events who are also members of the IWPT Programme Committee. Joakim Nivre makes a connection with learning dependency grammars in the CONLL-07 shared task, and the organisers of the Deep Linguistic Processing workshop discuss the ways in which broad coverage parsing systems can be developed for linguistically expressive grammars.

I would to thank all the programme committee members for their careful and timely work, especially those that took up extra rewiewing obligations at very short notice. Special thanks go to Paola Merlo, the programme chair, for organising the reviewing, designing the workshop programme and producing the proceedings. The scientific programme includes 14 full papaer and 3 short papers out of 31 submissions (of which 6 short papers). They cover all topics in parsing, from efficiency issues and complextiy of algorithms to accurate supervised and unsupervised learning techniques for parsing.

Harry Bunt
IWPT 2007 General Chair

Organizers

General Chair:

Harry Bunt (Tilburg University, Netherlands)

Programme Chair:

Paola Merlo (University of Geneva, Switzerland)

Logistic Arrangements Chair:

Alon Lavie (Carnegie-Mellon University, Pittsburgh, USA)

Programme Committee:

Harry Bunt (Tilburg University, Netherlands)
David Chiang(USC/ISI,USA)
John Carroll (University of Sussex, Brighton, UK)
Stephen Clark (Oxford University, UK)
James Henderson (University of Edinburgh, UK)
Ulf Hermjakob (USC Information Sciences Institute, Marina del Rey, USA)
Julia Hockenmaier (University of Pennsylvania, USA)
Aravind Joshi (University of Pennsylvania, Philadelphia, USA)
Ronald Kaplan (Xerox Palo Alto Research Center, USA)
Martin Kay (Xerox Palo Alto Research Center, USA)
Sadao Kurohashi (University of Tokyo, Japan)
Alon Lavie (Carnegie-Mellon University, Pittsburgh, USA)
Rob Malouf (San Diego State University, USA)
Yuji Matsumoto (Nara Institute of Science and Technology, Japan)
Bob Moore (Microsoft, Redmond, USA)
Mark-Jan Nederhof (MPI, Groeningen, Netherlands)
Joakim Nivre (Vaxjo University, Sweden)
Gertjan van Noord (University of Groningen, Netherlands)
Stephan Oepen (University of Oslo, Norway)
Stefan Riezler (Xerox Palo Alto Research Center, USA)
Giorgio Satta (University of Padua, Italy)
Kenji Sagae (University of Tokyo, Japan)
Khalil Sima'an (University of Amsterdam, Netherlands)
Eric Villemonte de la Clergerie (INRIA, Rocquencourt, France)
K. Vijay-Shanker (University of Delaware, USA)
Dekai Wu (Hong Kong University of Science and Technology, China)

Invited Speaker:

Stuart Shieber, Harvard University

Co-located Event Spotlight Presenters:

Joakim Nivre (Vaxjo University, Sweden)
Organisers of the Deep Linguistic Processing Workshop

Table of Contents

Conference Program

Saturday, 23 June, 2007

9:00–9:35 Registration/Opening Remarks

9:35–10:10 *Using Self-Trained Bilexical Preferences to Improve Disambiguation Accuracy*
Gertjan van Noord

10:10–10:45 *Evaluating Impact of Re-training a Lexical Disambiguation Model on Domain Adaptation of an HPSG Parser*
Tadayoshi Hara, Yusuke Miyao and Jun'ichi Tsujii

Coffee break

11:15–11:50 *Semi-supervised Training of a Statistical Parser from Unlabeled Partially-bracketed Data*
Rebecca Watson, Ted Briscoe and John Carroll

11:50–12:05 *Adapting WSJ-Trained Parsers to the British National Corpus using In-Domain Self-Training*
Jennifer Foster, Joachim Wagner, Djam Seddah and Josef van Genabith

Co-located Event Spotlight Presentation

12:05–12:40 *The Impact of Deep Linguistic Processing on Parsing Technology*
Timothy Baldwin, Mark Dras, Julia Hockenmaier, Tracy Holloway King and Gertjan van Noord

Lunch break

14:00–14:35 *Improving the Efficiency of a Wide-Coverage CCG Parser*
Bojan Djordjevic, James Curran and Stephen Clark

14:35–15:10 *Efficiency in Unification-Based N-Best Parsing*
Yi Zhang, Stephan Oepen and John Carroll

15:10–15:45 *A log-linear model with an n-gram reference distribution for accurate HPSG parsing*
Takashi Ninomiya, Takuya Matsuzaki, Yusuke Miyao and Jun'ichi Tsujii

Coffee break

Saturday, 23 June, 2007 (continued)

16:15–16:50 *Ambiguity Resolution by Reordering Rules in Text Containing Errors*
Sylvana Sofkova Hashemi

16:50–17:05 *Nbest Dependency Parsing with linguistically rich models*
Xiaodong Shi

17:05–17:40 *Symbolic Preference Using Simple Scoring*
Paula Newman

Sunday, 24 June, 2007

9:15–9:35 Registration

Guest Speaker

9:30–10:45 *Synchronous Grammars and Transducers: Good News and Bad News*
Stuart Shieber

Coffee break

11:15–11:50 *Are Very Large Context-Free Grammars Tractable?*
Pierre Boullier and Benot Sagot

11:50–12:05 *Pomset mcfgs*
Michael Pan

12:05–12:40 *Modular and Efficient Top-Down Parsing for Ambiguous Left-Recursive Grammars*
Richard Frost, Rahmatullah Hafiz and Paul Callaghan

Lunch break

14:00–14:35 *On the Complexity of Non-Projective Data-Driven Dependency Parsing*
Ryan McDonald and Giorgio Satta

14:35–15:10 *Dependency Parsing with Second-Order Feature Maps and Annotated Semantic Information*
Massimiliano Ciaramita and Giuseppe Attardi

Sunday, 24 June, 2007 (continued)

15:10–15:45 *A Latent Variable Model for Generative Dependency Parsing*
Ivan Titov and James Henderson

Coffee break

16:15–16:50 *Three-Dimensional Parametrization for Parsing Morphologically Rich Languages*
Reut Tsarfaty and Khalil Sima'an

Co-located Event Spotlight Presentation

16:50–17:25 *Data-Driven Dependency Parsing across Languages and Domains: Perspectives from the CoNLL-2007 Shared task*
Joakim Nivre