

IWPT-05

Proceedings of the Ninth International Workshop on Parsing Technologies

9–10 October 2005
Vancouver, British Columbia, Canada

Production and Manufacturing by
Omnipress Inc.
Post Office Box 7214
Madison, WI 53707-7214

©2005 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
75 Paterson Street, Suite 9
New Brunswick, NJ 08901
USA
Tel: +1-732-342-9100
Fax: +1-732-342-9339
acl@aclweb.org

Preface

The 9th International Workshop on Parsing Technologies continues the tradition that was started with the first workshop with that name, organized by Masaru Tomita in 1989 in Pittsburgh. IWPT'89 was followed by seven workshops in a biennial rhythm, only slightly disturbed by fear of millenium problems in 1999:

- IWPT'91 in Cancun, Mexico
- IWPT'93 in Tilburg, The Netherlands
- IWPT'95 in Prague, Czech Republic
- IWPT'97 in Cambridge (MA), USA
- IWPT 2000 in Trento, Italy
- IWPT 2001 in Beijing, China
- IWPT 2003 in Nancy, France

Over the years, the IWPT workshops have become the major forum for researchers in natural language parsing to meet and discuss advances in the field. Based on these workshops, four books on parsing technologies have been published by Kluwer Academic Publishers. The most recent one, based on IWPT 2000 and IWPT'01 was published last year (*New Developments in Parsing Technology*, Harry Bunt, John Carroll and Giorgio Satta, editors).

In 1994 the Special Interest Group on Parsing (SIGPARSE) was set up within ACL with the primary aim to give continuity to the IWPT series, and ever since the IWPT workshops have taken place under the auspices of SIGPARSE. IWPT 2005 also marks the return of IWPT to North America; the combined conferences on Human Language Technology and Empirical Methods in Natural Language Processing providing an excellent opportunity for this. I would like to thank Alon Lavie and the ACL organisation, in particular Priscilla Rasmussen, for their wonderful and professional help in organizing IWPT 2005 as a satellite event of the joint HLT/EMNLP conference.

Thanks are also due to the members of the Program Committee, and in particular to chairman Rob Malouf, for the careful work in reviewing submitted papers and designing the workshop program. Of a total of 42 submitted papers, 17 were accepted for presentation as a full paper, which gives an acceptance rate of 40%.

Harry Bunt
SIGPARSE Officer and IWPT 2005 General Chair

General Chair:

Harry Bunt (Tilburg University, Netherlands)

Program Chair:

Robert Malouf (San Diego State University, USA)

Logistic Arrangements Chair:

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

Program Committee:

Harry Bunt (Tilburg University, Netherlands)
Bob Carpenter (Alias-i, Inc., Brooklyn, NY, USA)
John Carroll (University of Sussex, Brighton, UK)
Eugene Charniak (Brown University, Providence, RI, USA)
Aravind Joshi (University of Pennsylvania, Philadelphia, USA)
Ronald Kaplan (Xerox Palo Alto Research Center, USA)
Martin Kay (Xerox Palo Alto Research Center, USA)
Dan Klein (UC Berkeley, USA)
Sadao Kurohashi (University of Tokyo, Japan)
Alon Lavie (Carnegie-Mellon University, Pittsburgh, USA)
Robert Malouf (San Diego State University, USA)
Yuji Matsumoto (Nara Institute of Science and Technology, Japan)
Paola Merlo (University of Geneva, Switzerland)
Bob Moore (Microsoft, Redmond, USA)
Anton Nijholt (University of Twente, Netherlands)
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)
Khalil Sima'an (University of Amsterdam, Netherlands)
Mark Steedman (University of Edinburgh, UK)
Hozumi Tanaka (Chukyo University, Japan)
Hans Uszkoreit (DFKI and Univeritat des Saarlandes, Germany)
Eric Villemonte de la Clergerie (INRIA, Rocquencourt, France)
K. Vijay-Shanker (University of Delaware, Newark, USA)
Dekai Wu (Hong Kong University of Science and Technology, China)

Additional Reviewers:

Rieks op de Akker (Twente University, Netherlands)

Daisuke Kawahara (University of Tokyo, Japan)
Manabu Okumura (Tokyo Institute of Technology, Japan)
Kenji Sagae (Carnegie-Mellon University, USA)
Libin Shen (University of Pennsylvania, USA)
Kiyooki Shirai (Hokuriku Advanced Institute of Science and Technology, Japan)

Invited Speakers:

Jill Burstein (Educational Testing Service, USA)
Mari Ostendorf (University of Washington, USA)

Table of Contents

<i>Efficient and Robust LFG Parsing: SxLFG</i>	
Pierre Boullier and Benoît Sagot	1
<i>Parsing Linear Context-Free Rewriting Systems</i>	
Håkan Burden and Peter Ljunglöf	11
<i>Switch Graphs for Parsing Type Logical Grammars</i>	
Bob Carpenter and Glyn Morrill	18
<i>Parsing with Soft and Hard Constraints on Dependency Length</i>	
Jason Eisner and Noah A. Smith	30
<i>Corrective Modeling for Non-Projective Dependency Parsing</i>	
Keith Hall and Václav Novák	42
<i>Better k-best Parsing</i>	
Liang Huang and David Chiang	53
<i>Machine Translation as Lexicalized Parsing with Hooks</i>	
Liang Huang, Hao Zhang and Daniel Gildea	65
<i>Treebank transfer</i>	
Martin Jansche	74
<i>Lexical and Structural Biases for Function Parsing</i>	
Gabriele Musillo and Paola Merlo	83
<i>Probabilistic Models for Disambiguation of an HPSG-Based Chart Generator</i>	
Hiroko Nakanishi, Yusuke Miyao and Jun'ichi Tsujii	93
<i>Efficacy of Beam Thresholding, Unification Filtering and Hybrid Parsing in Probabilistic HPSG Parsing</i>	
Takashi Ninomiya, Yoshimasa Tsuruoka, Yusuke Miyao and Jun'ichi Tsujii	103
<i>Head-Driven PCFGs with Latent-Head Statistics</i>	
Detlef Prescher	115
<i>A Classifier-Based Parser with Linear Run-Time Complexity</i>	
Kenji Sagae and Alon Lavie	125
<i>Chunk Parsing Revisited</i>	
Yoshimasa Tsuruoka and Jun'ichi Tsujii	133
<i>Constituent Parsing by Classification</i>	
Joseph Turian and I. Dan Melamed	141

<i>Strictly Lexical Dependency Parsing</i>	
Qin Iris Wang, Dale Schuurmans and Dekang Lin	152
<i>Efficient Extraction of Grammatical Relations</i>	
Rebecca Watson, John Carroll and Ted Briscoe	160
<i>Improving Parsing Accuracy by Combining Diverse Dependency Parsers</i>	
Daniel Zeman and Zdeněk Žabokrtský	171
<i>Exploring Features for Identifying Edited Regions in Disfluent Sentences</i>	
Qi Zhang and Fuliang Weng	179

Short papers

<i>Statistical Shallow Semantic Parsing despite Little Training Data</i>	
Rahul Bhagat, Anton Leuski and Eduard Hovy	186
<i>The Quick Check Pre-unification Filter for Typed Grammars: Extensions</i>	
Liviu Ciortuz	188
<i>From metagrammars to factorized TAG/TIG parsers</i>	
Éric Villemonte de la Clergerie	190
<i>Parsing Generalized ID/LP Grammars</i>	
Michael W. Daniels	192
<i>TFLEX: Speeding Up Deep Parsing with Strategic Pruning</i>	
Myroslava O. Dzikovska and Carolyn P. Rose	194
<i>Generic Parsing for Multi-Domain Semantic Interpretation</i>	
Myroslava Dzikovska, Mary Swift, James Allen and William de Beaumont	196
<i>Online Statistics for a Unification-Based Dialogue Parser</i>	
Micha Elsner, Mary Swift, James Allen and Daniel Gildea	198
<i>SUPPLE: A Practical Parser for Natural Language Engineering Applications</i>	
Robert Gaizauskas, Mark Hepple, Horacio Saggion, Mark A. Greenwood and Kevin Humphreys	200
<i>k-NN for Local Probability Estimation in Generative Parsing Models</i>	
Deirdre Hogan	202
<i>Robust Extraction of Subcategorization Data from Spoken Language</i>	
Jianguo Li, Chris Brew and Eric Fosler-Lussier	204