IWPT 2011

Proceedings of the 12th International Conference on Parsing Technologies

October 5-7, 2011
Dublin City University

© 2011 The Association for Computational Linguistics

Association for Computational Linguistics (ACL)
75 Paterson Street, Suite 9
New Brunswick, NJ 08901
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-732-342-9100
Fax: +1-732-342-9339
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-932432-04-6

Preface

Welcome to the 12th International Conference on Parsing Technologies (IWPT 2011) in Dublin, Ireland. IWPT 2011 continues the tradition of biennial conferences organized by SIGPARSE, ACL's Special Interest Group on Parsing, serving as the primary specialized forum for research on natural language parsing.

This year we received a total of 64 valid submissions, 42 long papers and 22 short papers, 6 of which were later withdrawn after being accepted for publication elsewhere. Of the remaining 58 submissions, 28 were accepted for presentation at the conference, which gives an acceptance rate of 48%. After notification, 2 more papers were withdrawn, which brings the final number of accepted papers to 26, all of which are published in these proceedings and presented at the conference in one of two ways: (i) as a long talk (long papers only) or (ii) as a short talk and a poster (short papers and some long papers). In this way, we were able to accommodate as many papers as possible and still give all the authors the opportunity of an oral presentation.

In addition to the contributed papers, IWPT 2011 will as usual feature invited talks on topics relevant to natural language parsing. This year we are delighted to welcome three very distinguished researchers: Ina Bornkessel-Schlesewsky, Michael Collins, and Mark Steedman. You will find the abstracts of their talks in the proceedings. There will also be a special workshop devoted to parsing of morphologically rich languages on the second day of the conference, a workshop that has had its own program committee and selection process.

Organizing IWPT 2011 would not have been possible without the dedicated work of a number of people. First and foremost, we would like to thank the local organizing committee, chaired by Özlem Çetinoğlu, who has done an outstanding job in taking care of all the local and practical organization. We are also grateful to the members of the program committee, who worked hard to review and discuss papers in the middle of the holiday season, and to the staff of SoftConf for support in managing the START system. Finally, thanks are due to the sponsors whose support helped to make IWPT 2011 possible: the Science Foundation Ireland, Dublin City University, the Centre for Next Generation Localisation, Springer Publishers, and Fáilte Ireland.

Enjoy the conference!

Harry Bunt	Joakim Nivre
General Chair	Program Chair

Organizers

General Chair:

Harry Bunt, Tilburg University

Program Chair:

Joakim Nivre, Uppsala University

Local Arrangements Chair:

Özlem Çetinoğlu, Dublin City University

Program Committee:

Timothy Baldwin, University of Melbourne
Srinivas Bangalore, AT&T
Philippe Blache, CNRS/Provence University
Harry Bunt, Tilburg University
Aoife Cahill, University of Stuttgart
John Carroll, University of Sussex
Eugene Charniak, Brown University
David Chiang, USC Information Sciences Institute
Stephen Clark, Cambridge University
Eric Villemonte de la Clergerie, INRIA Rocquencourt
Michael Collins, Columbia University
Özlem Çetinoğlu, Dublin City University
Jennifer Foster, Dublin City University
Josef van Genabith, Dublin City University
Carlos Gómez-Rodríguez, University of La Coruña
James Henderson, University of Geneva
Julia Hockenmaier, University of Illinois at Urbana-Champaign
Liang Huang, University of Southern California
Wenbin Jiang, Chinese Academy of Sciences
Mark Johnson, Macquarie University
Ronald Kaplan, Microsoft Corp
Martin Kay, Stanford University
Terry Koo, Google Inc
Sandra Kübler, Indiana University
Marco Kuhlmann, Uppsala University
Jonas Kuhn, University of Stuttgart
Sadao Kurohashi, Kyoto University
Alon Lavie, Carnegie Mellon University
Yuji Matsumoto, Nara Institute of Science and Technology
David McClosky, Stanford University
Ryan McDonald, Google Inc
Paola Merlo, University of Geneva
Yusuke Miyao, University of Tokyo
Mark-Jan Nederhof, University of St. Andrews
Joakim Nivre, University of Uppsala (chair)

Stephan Oepen, University of Oslo
Lilja Øvrelid, University of Oslo
Gerald Penn, University of Toronto
Slav Petrov, Google Inc
Owen Rambow, Columbia University
Kenji Sagae, Institute for Creative Technologies
Vijay Shanker, University of Delaware
Khalil Sima'an, University of Amsterdam
David Smith, University of Massachusetts
Noah Smith, Carnegie Mellon University
Mark Steedman, University of Edinburgh
Ivan Titov, Saarland University
Reut Tsarfaty, Uppsala University
Shuly Wintner, University of Haifa
Dekai Wu, Hong Kong University of Science and Technology
Yi Zhang, Saarland University
Yue Zhang, Cambridge University

Invited Speakers:

Ina Bornkessel-Schlesewsky, Philipps-Universität Marburg
Michael Collins, Columbia University
Mark Steedman, University of Edinburgh

Table of Contents

Invited Talk: *Computing Scope in a CCG Parser*
Mark Steedman 1

A Generalized View on Parsing and Translation
Alexander Koller and Marco Kuhlmann 2

Tree Parsing with Synchronous Tree-Adjoining Grammars
Matthias Büchse, Mark-Jan Nederhof and Heiko Vogler 14

Finding the Most Probable String and the Consensus String: an Algorithmic Study
Colin de la Higuera and Jose Oncina 26

A Word Clustering Approach to Domain Adaptation: Effective Parsing of Biomedical Texts
Marie Candito, Enrique Henestroza Anguiano and Djamé Seddah 37

Sentence-Level Instance-Weighting for Graph-Based and Transition-Based Dependency Parsing
Anders Søgaard and Martin Haulrich 43

Analysis of the Difficulties in Chinese Deep Parsing
Kun Yu, Yusuke Miyao, Takuya Matsuzaki, Xiangli Wang and Junichi Tsujii 48

On the Role of Explicit Morphological Feature Representation in Syntactic Dependency Parsing for German
Wolfgang Seeker and Jonas Kuhn 58

Bayesian Network Automata for Modelling Unbounded Structures
James Henderson 63

Model-Theory of Property Grammars with Features
Denys Duchier, Thi-Bich-Hanh Dao and Yannick Parmentier 75

Learning Structural Dependencies of Words in the Zipfian Tail
Tejaswini Deoskar, Markos Mylonakis and Khalil Sima'an 80

One-Step Statistical Parsing of Hybrid Dependency-Constituency Syntactic Representations
Kais Dukes and Nizar Habash 92

PLCFRS Parsing of English Discontinuous Constituents
Kilian Evang and Laura Kallmeyer 104

Invited Talk: *Towards a Neurobiologically Plausible Model of Human Sentence Comprehension Across Languages*
Ina Bornkessel-Schlesewsky 117

Minimally Supervised Domain-Adaptive Parse Reranking for Relation Extraction
Feiyu Xu, Hong Li, Yi Zhang, Hans Uszkoreit and Sebastian Krause 118

Simple Semi-Supervised Learning for Prepositional Phrase Attachment
Gregory F. Coppola, Alexandra Birch, Tejaswini Deoskar and Mark Steedman 129

Active Learning for Dependency Parsing Using Partially Annotated Sentences
Seyed Abolghasem Mirroshandel and Alexis Nasr 140

Invited Talk: *Lagrangian Relaxation for Inference in Natural Language Processing*
Michael Collins 150

Prefix Probabilities for Linear Context-Free Rewriting Systems
Mark-Jan Nederhof and Giorgio Satta 151

Efficient Matrix-Encoded Grammars and Low Latency Parallelization Strategies for CYK
Aaron Dunlop, Nathan Bodenstab and Brian Roark 163

Efficient Parallel CKY Parsing on GPUs
Youngmin Yi, Chao-Yue Lai, Slav Petrov and Kurt Keutzer 175

CuteForce – Deep Deterministic HPSG Parsing
Gisle Ytrestøl 186

Large-Scale Corpus-Driven PCFG Approximation of an HPSG
Yi Zhang and Hans-Ulrich Krieger 198

Features for Phrase-Structure Reranking from Dependency Parses
Richàrd Farkas, Bernd Bohnet and Helmut Schmid 209

Comparing the Use of Edited and Unedited Text in Parser Self-Training
Jennifer Foster, Özlem Çetinoğlu, Joachim Wagner and Josef van Genabith 215

Beyond Chart Parsing: An Analytic Comparison of Dependency Chart Parsing Algorithms
Meixun Jin, Hwidong Na and Jong-Hyeok Lee 220

Parser Evaluation Using Elementary Dependency Matching
Rebecca Dridan and Stephan Oepen 225

Parsing of Partially Bracketed Structures for Parse Selection
Mark-Jan Nederhof and Ricardo Sánchez-Sáez 231

Detecting Dependency Parse Errors with Minimal Resources
Markus Dickinson and Amber Smith 241

Conference Program

Wednesday October 5, 2011

09:15-09:30 Opening

09:30-10:30 Invited Talk

Computing Scope in a CCG Parser
Mark Steedman

10:30-11:00 Break

11:00-12:30 Long Talks 1

A Generalized View on Parsing and Translation
Alexander Koller and Marco Kuhlmann

Tree Parsing with Synchronous Tree-Adjoining Grammars
Matthias Büchse, Mark-Jan Nederhof and Heiko Vogler

Finding the Most Probable String and the Consensus String: an Algorithmic Study
Colin de la Higuera and Jose Oncina

12:30-14:00 Lunch

14:00-15:30 Short Talks 1

A Word Clustering Approach to Domain Adaptation: Effective Parsing of Biomedical Texts
Marie Candito, Enrique Henestroza Anguiano and Djamé Seddah

Sentence-Level Instance-Weighting for Graph-Based and Transition-Based Dependency Parsing
Anders Søgaard and Martin Haulrich

Analysis of the Difficulties in Chinese Deep Parsing
Kun Yu, Yusuke Miyao, Takuya Matsuzaki, Xiangli Wang and Junichi Tsujii

On the Role of Explicit Morphological Feature Representation in Syntactic Dependency Parsing for German
Wolfgang Seeker and Jonas Kuhn

Bayesian Network Automata for Modelling Unbounded Structures
James Henderson

Model-Theory of Property Grammars with Features
Denys Duchier, Thi-Bich-Hanh Dao and Yannick Parmentier

Wednesday October 5, 2011 (continued)

15:30-16:30 Break and Poster Session

16:30-18:00 Long Talks 2

Learning Structural Dependencies of Words in the Zipfian Tail
Tejaswini Deoskar, Markos Mylonakis and Khalil Sima'an

One-Step Statistical Parsing of Hybrid Dependency-Constituency Syntactic Representations
Kais Dukes and Nizar Habash

PLCFRS Parsing of English Discontinuous Constituents
Kilian Evang and Laura Kallmeyer

Thursday October 6, 2011

09:30-10:30 Invited Talk

Towards a Neurobiologically Plausible Model of Human Sentence Comprehension Across Languages
Ina Bornkessel-Schlesewsky

10:30-11:00 Break

11:00-12:30 Long Talks 3

Minimally Supervised Domain-Adaptive Parse Reranking for Relation Extraction
Feiyu Xu, Hong Li, Yi Zhang, Hans Uszkoreit and Sebastian Krause

Simple Semi-Supervised Learning for Prepositional Phrase Attachment
Gregory F. Coppola, Alexandra Birch, Tejaswini Deoskar and Mark Steedman

Active Learning for Dependency Parsing Using Partially Annotated Sentences
Seyed Abolghasem Mirroshandel and Alexis Nasr

12:30-13:30 Lunch

13:30-17:30 SPMRL Workshop

Friday October 7, 2011

09:30-10:30 Invited Talk

Lagrangian Relaxation for Inference in Natural Language Processing
Michael Collins

10:30-11:00 Break

11:00-12:30 Long Talks 4

Prefix Probabilities for Linear Context-Free Rewriting Systems
Mark-Jan Nederhof and Giorgio Satta

Efficient Matrix-Encoded Grammars and Low Latency Parallelization Strategies for CYK
Aaron Dunlop, Nathan Bodenstab and Brian Roark

Efficient Parallel CKY Parsing on GPUs
Youngmin Yi, Chao-Yue Lai, Slav Petrov and Kurt Keutzer

12:30-14:00 Lunch

14:00-15:30 Short Talks 2

CuteForce – Deep Deterministic HPSG Parsing
Gisle Ytrestøl

Large-Scale Corpus-Driven PCFG Approximation of an HPSG
Yi Zhang and Hans-Ulrich Krieger

Features for Phrase-Structure Reranking from Dependency Parses
Richàrd Farkas, Bernd Bohnet and Helmut Schmid

Comparing the Use of Edited and Unedited Text in Parser Self-Training
Jennifer Foster, Özlem Çetinoğlu, Joachim Wagner and Josef van Genabith

Beyond Chart Parsing: An Analytic Comparison of Dependency Chart Parsing Algorithms
Meixun Jin, Hwidong Na and Jong-Hyeok Lee

Parser Evaluation Using Elementary Dependency Matching
Rebecca Dridan and Stephan Oepen

15:30-16:30 Break and Poster Session

Friday October 7, 2011 (continued)

16:30-17:30 Long Talks 5

Parsing of Partially Bracketed Structures for Parse Selection
Mark-Jan Nederhof and Ricardo Sánchez-Sáez

Detecting Dependency Parse Errors with Minimal Resources
Markus Dickinson and Amber Smith

17:30-17:45 Closing