Introduction

The GEAF workshop aims to bring together grammar engineers from different frameworks to compare research and methodologies, particularly around the themes of evaluation, modularity, maintainability, relevance to theoretical and computational linguistics, and applications of “deep” grammars to real-world domains and NLP tasks.

Recent years have seen the development of techniques and resources to support robust, deep grammatical analysis of natural language in real-world domains and applications. The demands of these types of tasks have resulted in significant advances in areas such as parser efficiency, hybrid statistical/symbolic approaches to disambiguation, and the acquisition of large-scale lexicons. The effective acquisition, development, maintenance and enhancement of grammars is a central issue in such efforts, and the size and complexity of realistic grammars makes these tasks extremely challenging; indeed, these tasks are often tackled in ways that have much in common with software engineering. This workshop aims to bring together grammar engineers from different frameworks — for example LFG, HPSG, TAG, CCG, dependency grammar — to compare their research and methodologies.

We would like to thank the program committee Jason Baldridge, Emily Bender, Miriam Butt, Aoife Cahill, John Carroll, Ann Copestake, Berthold Crysmann, Mary Dalrymple, Stefanie Dipper, Dan Flickinger, Josef van Genabith, Ron Kaplan, Montserrat Marimon, Yusuke Miyao, Owen Rambow, and Jesse Tseng for their detailed reviews of all of the submitted papers, and Prof. Jun’ichi Tsujii for agreeing to give the invited talk.

In addition, we would like to thank Mark Stevenson for helping with general COLING workshop organization and Roger Evans for helping create these proceedings.
Organizers:
Stephen Clark, Oxford University
Tracy Holloway King, Palo Alto Research Center

Programme Committee:
Jason Baldridge, University of Texas at Austin
Emily Bender, University of Washington
Miriam Butt, Universitat Konstanz
Aoife Cahill, Universitat Stuttgart
John Carroll, University of Sussex
Ann Copestake, Cambridge University
Berthold Crismann, Bonn
Mary Dalrymple, Oxford University
Stefanie Dipper, Ruhr-Universitat Bochum
Dan Flickinger, Stanford University
Josef van Genabith, Dublin City University
Ron Kaplan, Powerset
Montserrat Marimon, Universitat Pompeu Fabra
Yusuke Miyao, University of Tokyo
Owen Rambow, Columbia University
Jesse Tseng, CNRS

Invited Speaker:
Jun’ichi Tsujii, University of Tokyo and University of Manchester
# Table of Contents

**TuLiPA: Towards a Multi-Formalism Parsing Environment for Grammar Engineering**  
Laura Kallmeyer, Timm Lichte, Wolfgang Maier, Yannick Parmentier, Johannes Dellert and Kilian Evang  
1

**Making Speech Look Like Text in the Regulus Development Environment**  
Elisabeth Kron, Manny Rayner, Marianne Santaholma, Pierrette Bouillon and Agnes Lisowska  
9

**A More Precise Analysis of Punctuation for Broad-Coverage Surface Realization with CCG**  
Michael White and Rajakrishnan Rajkumar  
17

**Multilingual Grammar Resources in Multilingual Application Development**  
Marianne Santaholma  
25

**Speeding up LFG Parsing Using C-Structure Pruning**  
Aoife Cahill, John T. Maxwell III, Paul Meurer, Christian Rohrer and Victoria Rosén  
33

**From Grammar-Independent Construction Enumeration to Lexical Types in Computational Grammars**  
Lars Hellan  
41

**Designing Testsuites for Grammar-based Systems in Applications**  
Valeria de Paiva and Tracy Holloway King  
49

**Towards Domain-Independent Deep Linguistic Processing: Ensuring Portability and Re-Usability of Lexicalised Grammars**  
Kostadin Cholakov, Valia Kordoni and Yi Zhang  
57
Conference Programme

Sunday, August 24, 2008

9:00–9:15 Opening Remarks

9:15–10:30 Invited Talk by Jun’ichi Tsujii

10:30–11:00 Break

11:00–11:30 TuLiPA: Towards a Multi-Formalism Parsing Environment for Grammar Engineering
Laura Kallmeyer, Timm Lichte, Wolfgang Maier, Yannick Parmentier, Johannes Dellert and Kilian Evang

11:30–12:00 Making Speech Look Like Text in the Regulus Development Environment
Elisabeth Kron, Manny Rayner, Marianne Santaholma, Pierrette Bouillon and Agnes Lisowska

12:00–12:30 A More Precise Analysis of Punctuation for Broad-Coverage Surface Realization with CCG
Michael White and Rajakrishnan Rajkumar

12:30–14:00 Lunch

14:00–14:30 Multilingual Grammar Resources in Multilingual Application Development
Marianne Santaholma

14:30–15:00 Speeding up LFG Parsing Using C-Structure Pruning
Aoife Cahill, John T. Maxwell III, Paul Meurer, Christian Rohrer and Victoria Rosén

15:00–15:30 From Grammar-Independent Construction Enumeration to Lexical Types in Computational Grammars
Lars Hellan

15:30–16:30 Demo Session with break

TuLiPA: Towards a Multi-Formalism Parsing Environment for Grammar Engineering (Laura Kallmeyer, Timm Lichte, Wolfgang Maier, Yannick Parmentier, Johannes Dellert and Kilian Evang)

Natural Language Entailment Using Implicit Information: An XLE Implementation (Daniel G. Bobrow, Bob Cheslow, Cleo Condoravdi, Lauri Karttunen, Tracy Holloway King, Charlotte Price and Annie Zaenen)

The Regulus Development Environment (Elisabeth Kron, Manny Rayner, Pierrette Bouillon, Marianne Santaholma and Agnes Lisowska)
Sunday, August 24, 2008 (continued)

MedSLT: Rule-based Medical Speech Translation System (Pierrette Bouillon, Manny Rayner, Sonia Halimi, Beth Ann Hockey, Hitoshi Isahara, Kyoko Kanzaki, Yukie Nakao, Marianne Santaholma, Marianne Starlander and Nikos Tsourakis)

Grammar and Output Representations in the C&C CCG Parser (Laura Rimell and Stephen Clark)

Developing a Modular Parsing System for Semantic Analysis of Japanese: the Verb Phrase Module (Yukiko Sasaki Alam)

The Checkpoint System: Hybrid Processing for Grammar and Style Checking (Tina Klüwer, Peter Adolphs and Berthold Crysmann)

Cognitive Grammar-based Linguistic Processor for Knowledge Extraction from Russian and English Texts (Igor Kuznetsov and Elena Kozerenko)

Defining and Viewing a Cross-linguistic Ontology of Verb Constructions (Lars Helan)

16:30–17:00  Designing Testsuites for Grammar-based Systems in Applications
Valeria de Paiva and Tracy Holloway King

17:00–17:30  Towards Domain-Independent Deep Linguistic Processing: Ensuring Portability and Re-Usability of Lexicalised Grammars
Kostadin Cholakov, Valia Kordoni and Yi Zhang