

**Proceedings of the  
13th International Conference on  
Finite-State Methods and Natural Language Processing  
FSMNLP 2017  
Umeå, Sweden**

**Edited by  
Frank Drewes**

## Introduction

These are the proceedings of the *13th International Conference on Finite State Methods and Natural Language Processing* (FSMNLP 2013), which was held from September 4 to September 6, 2017 in Umeå, Sweden, co-located with the *13th International Workshop on Tree Adjoining Grammars and Related Formalisms* (TAG+13), in order to foster the scientific exchange between the two research communities.

The conference series FSMNLP is the premier forum of the ACL Special Interest Group on Finite-State Methods (SIGFSM). It serves researchers and practitioners working on natural language processing (NLP) applications or language resources, theoretical and implementational aspects, or their combinations, that make use of finite-state methods.

FSMNLP 2017 received 17 submissions, each of which was carefully reviewed by at least three experts in the field. On the basis of these reviews the program committee selected eleven papers to be accepted for presentation at the meeting. Unfortunately, one of these papers had to be withdrawn later for budgetary reasons, which left ten papers for presentation and inclusion in this volume.

In addition to the regular papers, two invited talks were presented and three tutorials given:

**Invited talk 1** Gregory Kobele.

*Higher Order Structures in Minimalist Derivations*

**Invited talk 2** Ann Copestake.

*Dependency Semantics and Composition*

**Tutorial 1** Johanna Björklund.

*Minimization Techniques for Automata and Grammars*

**Tutorial 2** Sylvain Pogodalla.

*Abstract Categorical Grammars as a Model of the Syntax-Semantics Interface for TAG*

**Tutorial 3** Laura Kallmeyer and Rainer Osswald.

*Syntax-Driven Semantic Frame Composition in Lexicalized Tree Adjoining Grammars*

I would like to express my deepest gratitude to all authors for submitting their papers to FSMNLP 2017, to the members of the program committee for their excellent work in selecting the best papers, and to the members of the research group *Foundations of Language Processing* at Umeå University for their help with the local arrangements.

Frank Drewes  
chair

# Organization

## Program Chair

Frank Drewes

## Invited Speakers

Ann Copestake, University of Cambridge

Gregory Kobele, University of Chicago and University of Leipzig

## Tutorial Speakers

Johanna Björklund, Umeå University

Laura Kallmeyer & Rainer Osswald, Heinrich-Heine-Universität Düsseldorf

Sylvain Pogodalla, INRIA Nancy

## Program Committee

Matthieu Constant, Université de Lorraine

Frank Drewes, Umeå University (chair)

Tim Fernando, Trinity College Dublin

Thomas Hanneforth, University of Potsdam

Colin de La Higuera, Nantes University

Måns Hulden, University of Colorado Boulder

Laura Kallmeyer, Heinrich-Heine-Universität Düsseldorf

Kevin Knight, University of Southern California/ISI

András Kornai, Budapest Institute of Technology & Hungarian Academy of Sciences

Marco Kuhlmann, Linköping University

Andreas Maletti, University of Leipzig

Mark-Jan Nederhof, University of St Andrews

Jakub Piskorski, Polish Academy of Sciences

Laurette Pretorius, University of South Africa, Pretoria

Giorgio Satta, University of Padua

Tatjana Scheffler, University of Potsdam

Heiko Vogler, TU Dresden

Bruce Watson, Stellenbosch University

Anssi Yli-Jyrä, University of Helsinki

Menno van Zaanen, Tilburg University

Lynette van Zijl, Stellenbosch University

## Local Arrangements and Support

Suna Bensch

Henrik Björklund

Johanna Björklund

Petter Ericson

Anna Jonsson

Michael Minock

## Table of Contents

<i>Failure Transducers and Applications in Knowledge-Based Text Processing</i> Stoyan Mihov and Klaus U. Schulz .....	1
<i>Transliterated Mobile Keyboard Input via Weighted Finite-State Transducers</i> Lars Hellsten, Brian Roark, Prasoon Goyal, Cyril Allauzen, Françoise Beaufays, Tom Ouyang, Michael Riley and David Rybach .....	10
<i>Harmonic Serialism and Finite-State Optimality Theory</i> Yiding Hao .....	20
<i>Bounded-Depth High-Coverage Search Space for Noncrossing Parses</i> Anssi Yli-Jyrä .....	30
<i>Multi-tape Computing with Synchronous Relations</i> Christian Wurm and Simon Petitjean .....	41
<i>Finite-State Morphological Analysis for Marathi</i> Vinit Ravishankar and Francis M. Tyers .....	50
<i>Word Transduction for Addressing the OOV Problem in Machine Translation for Similar Resource-Scarce Languages</i> Shashikant Sharma and Anil Kumar Singh .....	56
<i>A FST Description of Noun and Verb Morphology of Azarbaijani Turkish</i> Razieh Ehsani, Berke Özenç and Ercan Solak .....	64
<i>Evaluation of Finite State Morphological Analyzers Based on Paradigm Extraction from Wiktionary</i> Ling Liu and Mans Hulden .....	69
<i>Evaluating an Automata Approach to Query Containment</i> Michael Minock .....	75

# Conference Program

## Monday, September 4, 2017

8:30–9:00 *Registration and Opening*

9:00–10:30 *Minimization Techniques for Automata and Grammars*  
Tutorial by Johanna Björklund

### 10:30–11:00 *Coffee Break*

11:00–12:30 *Abstract Categorical Grammars as a Model of the Syntax-Semantics Interface for TAG*  
Tutorial by Sylvain Pogodalla

### 12:30–13:30 *Lunch Break*

13:30–15:00 *Syntax-Driven Semantic Frame Composition in Lexicalized Tree Adjoining Grammars*  
Tutorial by Laura Kallmeyer and Rainer Osswald

### 15:00–15:30 *Coffee Break*

15:30–16:00 *Failure Transducers and Applications in Knowledge-Based Text Processing*  
Stoyan Mihov and Klaus U. Schulz

16:00–16:30 *Transliterated Mobile Keyboard Input via Weighted Finite-State Transducers*  
Lars Hellsten, Brian Roark, Prasoon Goyal, Cyril Allauzen, Françoise Beaufays, Tom Ouyang, Michael Riley and David Rybach

## Tuesday, September 5, 2017

9:00–10:00 *Higher Order Structures in Minimalist Derivations*  
Invited talk by Gregory Kobele

### 10:00–10:30 *Coffee Break*

10:30–11:00 *Harmonic Serialism and Finite-State Optimality Theory*  
Yiding Hao

11:00–11:30 *Bounded-Depth High-Coverage Search Space for Noncrossing Parses*  
Anssi Yli-Jyrä

11:30–12:00 *Multi-tape Computing with Synchronous Relations*  
Christian Wurm and Simon Petitjean

### 12:00–13:00 *Lunch Break*

## **Tuesday, September 5, 2017 (continued)**

- 13:00–13:30 *Finite-State Morphological Analysis for Marathi*  
Vinit Ravishankar and Francis M. Tyers
- 13:30–14:00 *Word Transduction for Addressing the OOV Problem in Machine Translation for Similar Resource-Scarce Languages*  
Shashikant Sharma and Anil Kumar Singh
- 14:00–14:30 *A FST Description of Noun and Verb Morphology of Azarbaijani Turkish*  
Razieh Ehsani, Berke Özenç and Ercan Solak
- 14:30–15:00** *Coffee Break*
- 15:00–16:30 SIGFSM Business Meeting

## **Wednesday, September 6, 2017**

- 9:00–10:00 *Dependency Semantics and Composition*  
Invited talk by Ann Copestake
- 10:00–10:30** *Coffee Break*
- 10:30–11:00 *Evaluation of Finite State Morphological Analyzers Based on Paradigm Extraction from Wiktionary*  
Ling Liu and Mans Hulden
- 11:00–11:30 *Evaluating an Automata Approach to Query Containment*  
Michael Minock

## **Closing and Lunch**