FSMNLP 2019

The 14th International Conference on Finite-State Methods and Natural Language Processing

Proceedings of the Conference

September 23 - 25, 2019
Dresden, Germany

Introduction

These are the proceedings of the 14th International Conference on Finite-State Methods and Natural Language Processing (FSMNLP 2019), which was held September 23-25, 2019 in Dresden, Germany.

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 2019 received 20 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 12 papers to be accepted for presentation at the conference.

In addition to the regular papers, one tutorial and two invited talks were presented:

Tutorial:

Aarne Ranta (University of Gothenburg, Sweden)
Grammatical Framework: an Interlingual Grammar Formalism

Invited talks:

Frank Drewes (Umeå University, Sweden)
A Survey of Recent Advances in Efficient Parsing for Graph Grammars

Kilian Gebhardt (Technische Universität Dresden, Germany)
Latent Variable Grammars for Discontinuous Parsing

We would like to express our gratitude to all authors for submitting their papers to FSMNLP 2019, to the members of the program committee and the four subreviewers for their excellent work in selecting the best papers, and to the members of the research group Foundations of Programmming Languages, TU Dresden for their help with the local arrangements.

Andreas Maletti and Heiko Vogler
co-chairs

Program Co-Chairs

Andreas Maletti and Heiko Vogler

Tutorial Speaker

Aarne Ranta, University of Gothenburg, Sweden

Invited Speakers

Frank Drewes, Umeå University, Sweden

Kilian Gebhardt, Technische Universität Dresden, Germany

Program Committee

Johanna Björklund (Umeå Universitet, Sweden)

Mathieu Constant (Université de Lorraine, France)

Jan Daciuk (Gdańsk University of Technology, Poland)

Frank Drewes (Umeå Universitet, Sweden)

Tim Fernando (Trinity College Dublin, Ireland)

Mike Hammond (University of Arizona, USA)

Thomas Hanneforth (Universität Potsdam, Germany)

Colin de la Higuera (Université de Nantes, France)

Mans Hulden (University of Colorado Boulder, USA)

András Kornai (Budapest Institute of Technology, Hungary)

Marco Kuhlmann (Linköpings Universitet, Sweden)

Andreas Maletti (Universität Leipzig, Germany, co-chair)

Mark-Jan Nederhof (University of St Andrews, UK)

Jakub Piskorski (Polish Academy of Sciences, Poland)

Anoop Sarkar (Simon Fraser University, Canada)

Richard Sproat (Google, USA)

Heiko Vogler (Technische Universität Dresden, Germany, co-chair)

Bruce William Watson (Stellenbosch University, South Africa)

Anssi Yli-Jyrä (University of Helsinki, Finland)

Menno van Zaanen (Universiteit van Tilburg, Netherlands)

Subreviewers

Henning Fernau, Universität Trier, Germany

Zoltan Fülöp, University of Szeged, Hungary

Kilian Gebhardt, Technische Universität Dresden, Germany

Luisa Herrmann, Technische Universität Dresden, Germany

Local Organization

Kerstin Achtruth

Kilian Gebhardt

Richard Mörbitz

Thomas Ruprecht

Heiko Vogler

Table of Contents

Grammatical Framework: an Interlingual Grammar Formalism
Aarne Ranta 1

A Survey of Recent Advances in Efficient Parsing for Graph Grammars
Frank Drewes 3

Latent Variable Grammars for Discontinuous Parsing
Kilian Gebhardt 4

Bottom-Up Unranked Tree-to-Graph Transducers for Translation into Semantic Graphs
Johanna Björklund, Shay B. Cohen, Frank Drewes and Giorgio Satta 7

On the Compression of Lexicon Transducers
Marco Cognetta, Cyril Allauzen and Michael Riley 18

MSO with tests and reducts
Tim Fernando, David Woods and Carl Vogel 27

Finite State Transducer Calculus for Whole Word Morphology
Maciej Janicki 37

Weighted parsing for grammar-based language models
Richard Mörbitz and Heiko Vogler 46

Regular transductions with MCFG input syntax
Mark-Jan Nederhof and Heiko Vogler 56

A Syntactically Expressive Morphological Analyzer for Turkish
Adnan Ozturel, Tolga Kayadelen and Isin Demirsahin 65

Using Meta-Morph Rules to develop Morphological Analysers: A case study concerning Tamil
Kengatharaiyer Sarveswaran, Gihan Dias and Miriam Butt 76

Distilling weighted finite automata from arbitrary probabilistic models
Ananda Theertha Suresh, Brian Roark, Michael Riley and Vlad Schogol 87

Silent HMMs: Generalized Representation of Hidden Semi-Markov Models and Hierarchical HMMs
Kei Wakabayashi 98

Latin script keyboards for South Asian languages with finite-state normalization
Lawrence Wolf-Sonkin, Vlad Schogol, Brian Roark and Michael Riley 108

Transition-Based Coding and Formal Language Theory for Ordered Digraphs
Anssi Yli-Jyrä 118

Conference Program

Sunday, September 22nd	
19:00	welcome reception (restaurant at Weberplatz 3, Dresden)

Monday, September 23rd	
08:00–08:50	registration
08:50–09:00	opening
09:00–09:45	Tutorial (part 1) Aarne Ranta (University of Gothenburg, Sweden) *Grammatical Framework: an Interlingual Grammar Formalism*
09:45–10:15	coffee break
10:15–11:00	Tutorial (part 2) Aarne Ranta (University of Gothenburg, Sweden) *Grammatical Framework: an Interlingual Grammar Formalism*
11:00–11:30	Tutorial (part 3) Aarne Ranta (University of Gothenburg, Sweden) *Grammatical Framework: an Interlingual Grammar Formalism*
12:00–13:30	lunch break
13:30–14:00	Anssi Yli-Jyrä. *Transition-Based Coding and Formal Language Theory for Ordered Digraphs*
14:00–14:30	Marco Cognetta, Cyril Allauzen, and Michael Riley. *On the Compression of Lexicon Transducers*
14:30–15:00	coffee break
15:00–15:30	Ananda Theertha Suresh, Brian Roark, Michael Riley, and Vlad Schogol. *Distilling weighted finite automata from arbitrary probabilistic models*
15:30–16:00	Kei Wakabayashi. *Silent HMMs: Generalized Representation of Hidden Semi-Markov Models and Hierarchical HMMs*

Tuesday, September 24th	
09:00–10:00	Invited lecture: Frank Drewes (Umeå University, Sweden) *A Survey of Recent Advances in Efficient Parsing for Graph Grammars*
10:00–10:30	coffee break
10:30–11:00	Adnan Ozturel, Tolga Kayadelen, and Isin Demirsahin. *A Syntactically Expressive Morphological Analyzer for Turkish*
11:00–11:30	Kengatharaiyer Sarveswaran, Gihan Dias, and Mirriam Butt. *Using Meta-Morph Rules to develop Morphological Analysers: A case study concerning Tamil*
11:30–12:00	Maciej Janicki. *Finite State Transducer Calculus for Whole Word Morphology*
12:00–13:30	lunch break
13:30–14:00	Lawrence Wolf-Sonkin, Vlad Schogol, Brian Roark, and Michael Riley. *Latin script keyboards for South Asian languages with finite-state normalization*
14:00–14:30	Johanna Björklund, Shay B. Cohen, Frank Drewes, and Giorgio Satta. *Bottom-Up Unranked Tree-to-Graph Transducers for Translation into Semantic Graphs*
14:30–15:00	coffee break
15:00–16:30	SIGFSM business meeting
18:00–18:45	excursion (guided tour through Dresden)
19:00	conference dinner

Wednesday, September 25th	
09:00–10:00	Invited lecture: Kilian Gebhardt (Technische Universität Dresden, Germany) *Latent Variable Grammars for Discontinuous Parsing*
10:00–10:30	coffee break
10:30–11:00	Mark-Jan Nederhof and Heiko Vogler. *Regular transductions with MCFG input syntax*
11:00–11:30	Tim Fernando, David Woods, and Carl Vogel. *MSO with tests and reducts*
11:30–12:00	Richard Mörbitz and Heiko Vogler. *Weighted parsing for grammar-based language models*
12:00	closing of FSMNLP 2019