

EACL 2012

ATANLP 2012

**Workshop on
Applications of Tree Automata Techniques in
Natural Language Processing**

Proceedings of the Workshop

April 24 2012
Avignon France

© 2012 The Association for Computational Linguistics

ISBN 978-1-937284-19-0

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Preface

We are very pleased to present the proceedings of the EACL 2012 Workshop on Applications of Tree Automata Techniques in Natural Language Processing (ATANLP 2012), held on April 24 in Avignon, France. The first ATANLP workshop was held two years ago in Uppsala, Sweden, in connection with the 48th Annual Meeting of the Association for Computational Linguistics. This year, ATANLP is conducted as a workshop of the 13th Conference of the European Chapter of the Association for Computational Linguistics.

The theory of tree automata has always had a close connection with natural language processing. In the 1960s, computational linguistics was the major driving force for the development of a theory of tree automata. Nevertheless, the number of successful applications of this theory to natural language processing remained small during the 20th century. During the last decade, the situation has started to change, and the change accelerates. Applications of tree automata in natural language processing can be found in work on topics as diverse as grammar formalisms, computational semantics, language generation, and machine translation. Researchers in natural language processing have recognized the usefulness of tree automata theory for solving the problems they are interested in, and theorists are inspired by the resulting theoretical questions.

The goals of this workshop are to provide a dedicated venue for the presentation of work that relates the tree automata techniques to natural language processing, and to create a forum where researchers from the two areas can meet and exchange ideas. Specifically, the workshop aims at raising the awareness for theoretical results useful for applications in natural language processing, and at identifying open theoretical problems raised by such applications.

We are very happy that Andreas Maletti (University of Stuttgart, Germany), an expert in the area of weighted tree automata and their application to natural language processing problems, agreed to present the invited lecture of the workshop. In addition to the invited talk and the presentations of four regular papers, the workshop features two tutorials given by Alexander Koller and Heiko Vogler. The workshop is concluded by an open problem session that, as we hope and believe, will provide an inspiring outlook on possible venues for future research.

We thank the members of the program committee for their support, and in particular for being careful reviewers of the papers submitted. Furthermore, we would like to thank the program committee chairs, Mirella Lapata and Lluís Màrquez, as well as the workshop chairs, Kristiina Jokinen and Alessandro Moschitti, for their friendly and professional assistance.

We hope that all participants of the workshop experience an inspiring event characterized by curiosity and an open-minded atmosphere, and that all readers of these proceedings will gain new insights that make a difference.

Frank Drewes
Marco Kuhlmann
March 2012

Organizers:

Frank Drewes, Umeå University (Sweden)
Marco Kuhlmann, Uppsala University (Sweden)

Program Committee:

Parosh Aziz Abdulla, Uppsala University (Sweden)
Leonor Becerra-Bonache, Université Jean Monnet (France)
Johanna Björklund, Umeaa University (Sweden)
David Chiang, ISI/University of Southern California (USA)
Loek Cleophas, Eindhoven University of Technology (The Netherlands)
François Denis, LIF Marseille (France)
Laura Kallmeyer, University of Düsseldorf (Germany)
Kevin Knight, ISI/University of Southern California (USA)
Alexander Koller, University of Potsdam (Germany)
Sebastian Maneth, NICTA and University of New South Wales (Australia)
Brink van der Merwe, Stellenbosch University (South Africa)
Mark-Jan Nederhof, University of St Andrews (UK)
Joachim Niehren, INRIA Lille (France)
Kai Salomaa, Queen's University (Canada)
Marc Tommasi, INRIA Lille (France)
Heiko Vogler, Technische Universität Dresden (Germany)

Invited Speaker:

Andreas Maletti, University of Stuttgart (Germany)

Presenters of Tutorials:

Alexander Koller, University of Potsdam (Germany)
Heiko Vogler, TU Dresden (Germany)

Table of Contents

<i>Preservation of Recognizability for Weighted Linear Extended Top-Down Tree Transducers</i> Nina Seemann, Daniel Quernheim, Fabienne Braune and Andreas Maletti	1
<i>Deciding the Twins Property for Weighted Tree Automata over Extremal Semifields</i> Matthias Büchse and Anja Fischer	11
<i>TTT: A Tree Transduction Language for Syntactic and Semantic Processing</i> Adam Purtee and Lenhart Schubert	21
<i>Second Position Clitics and Monadic Second-Order Transduction</i> Neil Ashton	31

Conference Program

Tuesday, April 24, 2012

8:45-9:00 Opening

9:00–10:00 Invited Talk: Multi Bottom-up Tree Transducers by Andreas Maletti

10:00–10:30 Coffee Break

Paper Presentations

10:30–11:00 *Preservation of Recognizability for Weighted Linear Extended Top-Down Tree Transducers*
Nina Seemann, Daniel Quernheim, Fabienne Braune and Andreas Maletti

11:00–11:30 *Deciding the Twins Property for Weighted Tree Automata over Extremal Semifields*
Matthias Büchse and Anja Fischer

11:30–12:00 *TTT: A Tree Transduction Language for Syntactic and Semantic Processing*
Adam Purtee and Lenhart Schubert

12:00–12:30 *Second Position Clitics and Monadic Second-Order Transduction*
Neil Ashton

Lunch

14:30–15:30 Tutorial: Weighted Recognizable Tree Languages – A Survey by Heiko Vogler

15:30–16:00 Coffee Break

16:00–17:00 Tutorial: Interpreted Regular Tree Grammars by Alexander Koller

17:00–17:30 Open Problems

17:30 Closing

