

MoL 13

The 13th Meeting on the Mathematics of Language

Proceedings

August 9, 2013
Sofia, Bulgaria

Production and Manufacturing by
Omnipress, Inc.
2600 Anderson Street
Madison, WI 53704 USA

©2013 The Association for Computational Linguistics

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

ISBN 978-1-937284-65-7

Introduction

The Mathematics of Language (MoL) special interest group traces its origins to a meeting held in October 1984 at Ann Arbor, Michigan. While MoL is among the oldest SIGs of the ACL, it is the first time that the proceedings are produced by our parent organization. The first volume was published by Benjamins, later ones became special issues of the *Annals of Mathematics and Artificial Intelligence* and *Linguistics and Philosophy*, and for the last three occasions (really six years, since MoL only meets every second year) we relied on the Springer LNCS series. Perhaps the main reason for this aloofness was that the past three decades have brought the ascendancy of statistical methods in computational linguistics, with the formal, grammar-based methods that were the mainstay of mathematical linguistics viewed with increasing suspicion.

To make matters worse, the harsh anti-formal rhetoric of leading linguists relegated important attempts at formalizing Government-Binding and later Minimalist theory to the fringes of syntax. Were it not for phonology and morphology, where the incredibly efficient finite state methods pioneered by Kimmo Koskenniemi managed to bridge the gap between computational practice and linguistic theory, and were it not for the realization that the mathematical approach has no alternative in machine learning, MoL could have easily disappeared from the frontier of research.

The current volume marks a time when we can begin to see the computational and the theoretical linguistics camps together again. The selection of papers, while still strong on phonology (Heinz and Lai, Heinz and Rogers) and morphology (Kornai et al.), extends well to syntax (Hunter and Dyer, Fowlie) and semantics (Clark et al., Fernando). Direct computational concerns such as machine translation (Martzoukos et al.), decoding (Corlett and Penn), and complexity (Berglund et al.) are now clearly seen as belonging to the core focus of the field.

The 10 papers presented in this volume were selected by the Program Committee from 16 submissions. We would like to thank the authors, the members of the Program Committee, and our invited speaker for their contributions to the planning and execution of the workshop, and the ACL conference organizers, especially Aoife Cahill and Qun Liu (workshops), and Roberto Navigli and Jing-Shin Chang (publications) for their significant contributions to the overall management of the workshop and their direction in preparing the publication of the proceedings.

Andras Kornai and Marco Kuhlmann (editors)
June 2013

Program Chairs:

András Kornai (Hungarian Academy of Sciences, Hungary)

Marco Kuhlmann (Uppsala University, Sweden)

Program Committee:

Ash Asudeh (Carleton University, Canada)

Alexander Clark (King's College, UK)

Annie Foret (University of Rennes 1, France)

Daniel Gildea (University of Rochester, USA)

Gerhard Jäger (University of Tübingen, Germany)

Aravind Joshi (University of Pennsylvania, USA)

Makoto Kanazawa (National Institute of Informatics, Japan)

Greg Kobele (University of Chicago, USA)

Andreas Maletti (University of Stuttgart, Germany)

Carlos Martín-Vide (University Rovira i Virgili, Spain)

Jens Michaelis (Bielefeld University, Germany)

Gerald Penn (University of Toronto, Canada)

Carl Pollard (The Ohio State University, USA)

Jim Rogers (Earlham College, USA)

Giorgio Satta (University of Padua, Italy)

Noah Smith (Carnegie Mellon University, USA)

Ed Stabler (UCLA, USA)

Mark Steedman (Edinburgh University, UK)

Sylvain Salvati (INRIA, France)

Anssi Yli-Jyrä (University of Helsinki, Finland)

Invited Speaker:

Mark Johnson (Macquarie University, Australia)

Table of Contents

<i>Distributions on Minimalist Grammar Derivations</i> Tim Hunter and Chris Dyer	1
<i>Order and Optionality: Minimalist Grammars with Adjunction</i> Meaghan Fowlie	12
<i>On the Parameterized Complexity of Linear Context-Free Rewriting Systems</i> Martin Berglund, Henrik Björklund and Frank Drewes	21
<i>Segmenting Temporal Intervals for Tense and Aspect</i> Tim Fernando	30
<i>The Frobenius Anatomy of Relative Pronouns</i> Stephen Clark, Bob Coecke and Mehrnoosh Sadrzadeh	41
<i>Vowel Harmony and Subsequentiality</i> Jeffrey Heinz and Regine Lai	52
<i>Learning Subregular Classes of Languages with Factored Deterministic Automata</i> Jeffrey Heinz and Jim Rogers	64
<i>Structure Learning in Weighted Languages</i> Andras Kornai, Attila Zséder and Gábor Recski	72
<i>Why Letter Substitution Puzzles are Not Hard to Solve: A Case Study in Entropy and Probabilistic Search-Complexity</i> Eric Corlett and Gerald Penn	83
<i>Investigating Connectivity and Consistency Criteria for Phrase Pair Extraction in Statistical Machine Translation</i> Spyros Martzoukos, Christophe Costa Florêncio and Christof Monz	93
<i>Grammars and Topic Models</i> Mark Johnson	102

Program

Friday, August 9

Session 1

09:00–09:30 *Distributions on Minimalist Grammar Derivations*

Tim Hunter and Chris Dyer

09:30–10:00 *Order and Optionality: Minimalist Grammars with Adjunction*

Meaghan Fowlie

10:00–10:30 *On the Parameterized Complexity of Linear Context-Free Rewriting Systems*

Martin Berglund, Henrik Björklund and Frank Drewes

10:30–11:00 Coffee Break

Session 2

11:00–11:30 *Segmenting Temporal Intervals for Tense and Aspect*

Tim Fernando

11:30–12:00 *The Frobenius Anatomy of Relative Pronouns*

Stephen Clark, Bob Coecke and Mehrnoosh Sadrzadeh

12:00–12:30 *Vowel Harmony and Subsequentiality*

Jeffrey Heinz and Regine Lai

12:30–14:00 Lunch Break

Friday, August 9 (continued)

Session 3

- 14:00–14:30 *Learning Subregular Classes of Languages with Factored Deterministic Automata*
Jeffrey Heinz and Jim Rogers
- 14:30–15:00 *Structure Learning in Weighted Languages*
Andras Kornai, Attila Zséder and Gábor Recski
- 15:00–15:30 *Why Letter Substitution Puzzles are Not Hard to Solve: A Case Study in Entropy and Probabilistic Search-Complexity*
Eric Corlett and Gerald Penn
- 15:30–16:00 Coffee Break

Session 4

- 16:00–16:30 *Investigating Connectivity and Consistency Criteria for Phrase Pair Extraction in Statistical Machine Translation*
Spyros Martzoukos, Christophe Costa Florêncio and Christof Monz
- 16:30–17:30 *Grammars and Topic Models*
Mark Johnson