Welcome to Jeju Island — where ACL makes a return to Asia!

As General Chair, I am indeed honored to pen the first words of ACL 2012 proceedings. In the past year, research in computational linguistics has continued to thrive across Asia and all over the world. On this occasion, I share with you the excitement of our community as we gather again at our annual meeting. On behalf of the organizing team, it is my great pleasure to welcome you to Jeju Island and ACL 2012.

In 2012, ACL turns 50. I feel privileged to chair the conference that marks such an important milestone for our community. We have prepared special programs to commemorate the 50th anniversary, including ‘Rediscovering 50 Years of Discovery’, a main conference workshop chaired by Rafael Banchs with a program on ‘the People, the Contents, and the Anthology’, which recollects some of the great moments in ACL history, and ‘ACL 50th Anniversary Lectures’ by Mark Johnson, Aravind K. Joshi and a Lifetime Achievement Award Recipient.

A large number of people have worked hard to bring this annual meeting to fruition. It has been an unforgettable experience for everyone involved. My deepest thanks go to the authors, reviewers, volunteers, participants, and all members and chairs of the organizing committees. It is your participation that makes a difference.

Program Chairs, Chin-Yew Lin and Miles Osborne, deserve our gratitude for putting an immense amount of work to ensure that each of the 940 submissions was taken care of. They put together a superb technical program like nobody else. Publication Chairs, Maggie Li and Michael White, extended the publishing tools to take care of every detail and compiled all the books within an impossible schedule. Tutorial Chair, Michael Strube, put together six tutorials that you can never miss. Workshop Chairs, Massimo Poesio and Satoshi Sekine, working with their EACL and NAACL counterparts, selected 11 quality workshops, many of which are new editions in their popular workshop series. Demo Chair, Min Zhang, started a novel review process and selected 29 quality system demos. Faculty Advisors, Kentaro Inui, Greg Kondrak, and Yang Liu, and Student Chairs, Jackie Cheung, Jun Hatori, Carlos Henriquez and Ann Irvine, assembled an excellent program for the Student Research Workshop with 12 accepted papers. Mentoring Chair, Joyce Chai, coordinated the mentorship of 13 papers. Publicity Chairs, Jung-jae Kim and Youngjoong Ko, developed the website, newsletters, and conference handbook that kept us updated all the time. Exhibition Chair, Byeongchang Kim, coordinated more than 10 exhibitors with a strong industry presence. All the events are now brought to us on Jeju Island by the Local Arrangements Chairs, Gary Lee and Jong Park, and their team. I can never thank them enough for all the preparations they have made to host us in such a spectacular place!

I would like to express my gratitude and appreciation to Kevin Knight, Chair of the ACL Conference Coordination Committee, Dragomir Radev, ACL Secretary, and Priscilla Rasmussen, ACL Business Manager, for their advice and guidance throughout the process.

The financial sponsors generously supported ACL 2012 in a meaningful way despite a challenging
economic outlook. We are honored to have Baidu as the Platinum Sponsor, Elsevier and Google as Gold Sponsors, Microsoft, KAIST and SK as Silver Sponsors, 7 Bronze Sponsors, and 3 Supporters. The Donald and Betty Walker Student Scholarship Fund and Asian Federation of Natural Language Processing have supported our student travel grants. The sponsorship program was made possible by the ACL sponsorship committee: Eiichiro Sumita, Haifeng Wang, Michael Gamon, Patrick Pantel, Massimiliano Ciaramita, and Idan Szpektor.

Finally, I do hope that you have an enjoyable and productive time on Jeju Island, and that you will leave with fond memories of ACL’s 50th Anniversary. With my best wishes for a successful conference!

Haizhou Li
ACL 2012 General Chair
July 2012
Preface: Programme Committee Co-Chairs

This year we received 571 valid long paper submissions and 369 short paper submissions. 19% of the long papers and 20% of the short papers were accepted. As usual, some are presented orally and some as posters. Taking unigram counts from accepted long paper titles, and ignoring function words, the most popular word were:

entity 5
evaluation 5
hierarchical 5
information 5
joint 5
syntactic 5
topic 5
discriminative 6
lexical 6
statistical 6
chinese 7
dependency 7
machine 8
modeling 8
models 8
language 10
word 10
parsing 11
model 12
learning 14
translation 15

Some areas have grown over time and some have diminished. The most popular area for submissions (as expected) was Machine Translation. We promoted Social Media as a new area.

Twenty nine Area Chairs worked with 665 reviewers, producing 1830 long paper reviews and 1187 short paper reviews. Everything ran to a tight schedule and there were no slippages. This would not have been possible without our wonderful and diligent Area Chairs and Reviewers. Thanks!

We are delighted to have two keynote speakers, both of whom are very well known to the language community: Aravind Joshi and Mark Johnson. They will give coordinated talks addressing the 50th ACL anniversary: “Remembrance of ACLs past” and “Computational linguistics: Where do we go from here?” The ACL Lifetime Achievement Award will be announced on the last day of the conference.

Of the many papers, we selected two as being outstanding:
Bayesian Symbol-Refined Tree Substitution Grammars for Syntactic Parsing
Hiroyuki Shindo, Yusuke Miyao, Akinori Fujino, Masaaki Nagata

String Re-writing Kernel
Fan Bu, Hang Li, Xiaoyan Zhu

They will be presented as best papers in a dedicated session.

We thank the General Conference Chair Haizhou Li, the Local Arrangements Committee headed by Gary Geunbae Lee, Michael White and Maggie Li, the Publication Co-Chairs for coordinating and putting the proceedings together and all other committee chairs for their work. MO is especially thankful to Steve Clark for helpful tips on how to manage and run the whole process.

We hope you enjoy the conference!

Chin-Yew Lin, Microsoft Research Asia
Miles Osborne, University of Edinburgh
Organizing Committee

General Chair
Haizhou Li, Institute for Infocomm Research

Program Co-Chairs
Chin-Yew Lin, Microsoft Research Asia
Miles Osborne, University of Edinburgh

Local Arrangement Co-Chairs
Gary Geunbae Lee, Pohang University of Science and Technology (POSTECH)
Jong C. Park, Korea Advanced Institute of Science and Technology (KAIST)

Workshop Co-Chairs
Massimo Poesio, University of Essex
Satoshi Sekine, New York University

Publication Co-Chairs
Maggie Li, The Hong Kong Polytechnic University
Michael White, The Ohio State University

Publicity Chairs
Jung-jae Kim, Nanyang Technological University
Youngjoong Ko, Dong-A University

Tutorial Chair
Michael Strube, HITS gGmbH

Demo Chair
Min Zhang, Institute for Infocomm Research

Special Session Chair
Rafael E. Banchs, Institute for Infocomm Research

Mentoring Service Chair
Joyce Chai, Michigan State University

Exhibit Chair
Byeongchang Kim, Catholic University of Daegu

Faculty Advisors (Student Research Workshop)
Kentaro Inui, Tohoku University
Greg Kondrak, University of Alberta
Yang Liu, University of Texas at Dallas
Student Chairs (Student Research Workshop)

Jackie Cheung, University of Toronto
Jun Hatori, University of Tokyo
Carlos Henriquez, Technical University of Catalonia (UPC)
Ann Irvine, Johns Hopkins University

Sponsorship Chairs

Eiichiro Sumita, NICT
Haifeng Wang, Baidu
Michael Gamon, Microsoft
Patrick Pantel, Microsoft
Massimiliano Ciaramita, Google
Idan Szpektor, Yahoo!

Local Arrangements Committee

Gary Geunbae Lee (co-chair)
Jong C. Park (co-chair)
Jeong-Won Cha (social activities)
Hanmin Jung (local sponsorship)
Seunghik Kang (local finance)
Byeongchang Kim (local exhibit)
Harksoo Kim (government sponsorship)
Jung-jae Kim (conference handbook)
Youngjoong Ko (web site and flyer)
Heuiseok Lim (student volunteer management)
Seong-Bae Park (internet, wifi and equipments)

Business Manager

Priscilla Rasmussen
Program Committee

Program Co-chairs
Chin-Yew Lin, Microsoft Research Asia
Miles Osborne, University of Edinburgh

Area Chairs
Hongyuan Zha, School of Computational Science and Engineering, College of Computing, Georgia Institute of Technology
Hsin-Hsi Chen, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan
Kemal Oflazer, Carnegie Mellon University - Qatar
Mikio Nakano, Honda Research Institute, Japan
Andrei Popescu-Belis, Idiap Research Institute, Switzerland
Eneko Agirre, University of the Basque Country, Spain
Jason Baldridge, The University of Texas at Austin, USA
David Weir, University of Sussex, UK
Trevor Cohn, University of Sheffield, UK
Mark Dredze, Johns Hopkins University, USA
Noah Smith, Carnegie Mellon University, USA
Jan Wiebe, University of Pittsburgh, USA
Daniel M. Bikel, Google Research, USA
Mona T. Diab, Center for Computational Learning Systems, Columbia University, USA
Fei Xia, Univ. of Washington, Seattle, USA
Marcello Federico, Fondazione Bruno Kessler, Trento, Italy
Adam Lopez, Human Language Technology Center of Excellence, Johns Hopkins University, USA
David Talbot, Google Research, USA
Hua Wu, Baidu, China
Evgeniy Gabrilovich, Google, USA
Naoaki Okazaki, Graduate School of Information Sciences, Tohoku University, Japan
Stephen Clark, University of Cambridge, UK
Liang Huang, University Southern California, USA
Anja Belz, School of Computing, Engineering and Maths, University of Brighton, UK
Ani Nenkova, University of Pennsylvania, USA
Chung-Hsien Wu, Department of Computer Science and Information Engineering, National Cheng Kung University, TAIWAN
Jian Su, Institute for Infocomm Research, Singapore
Jianfeng Gao, Microsoft Research, Redmond, USA
Vincent Ng, University of Texas at Dallas, USA

Program Committee
Apoorv Agarwal, Lars Ahrenberg, Hua Ai, Cem Akkaya, Inaki Alegria, Jan Alexandersson, Enrique Alfonseca, Ben Allison, Sophia Ananiadou, Abhishek Arun
Table of Contents

Higher-order Constituent Parsing and Parser Combination  
Xiao Chen and Chunyu Kit ......................................................... 1

Joint Evaluation of Morphological Segmentation and Syntactic Parsing  
Reut Tsarfaty, Joakim Nivre and Evelina Andersson .......................... 6

A Comparison of Chinese Parsers for Stanford Dependencies  
Wanxiang Che, Valentin Spitkovsky and Ting Liu ............................ 11

A Feature-Rich Constituent Context Model for Grammar Induction  
Dave Golland, John DeNero and Jakob Uszkoreit ............................ 17

Private Access to Phrase Tables for Statistical Machine Translation  
Nicola Cancedda ........................................................................... 23

Fast and Scalable Decoding with Language Model Look-Ahead for Phrase-based Statistical Machine Translation  
Joern Wuebker, Hermann Ney and Richard Zens ............................... 28

Head-Driven Hierarchical Phrase-based Translation  
Junhui Li, Zhaopeng Tu, Guodong Zhou and Josef van Genabith ............ 33

Joint Learning of a Dual SMT System for Paraphrase Generation  
Hong Sun and Ming Zhou .................................................................. 38

A Novel Burst-based Text Representation Model for Scalable Event Detection  
Xin Zhao, Rishan Chen, Kai Fan, Hongfei Yan and Xiaoming Li ............ 43

A Graph-based Cross-lingual Projection Approach for Weakly Supervised Relation Extraction  
Seokhwan Kim and Gary Geunbae Lee ............................................. 48

Pattern Learning for Relation Extraction with a Hierarchical Topic Model  
Enrique Alfonseca, Katja Filippova, Jean-Yves Delort and Guillermo Garrido .... 54

Self-Disclosure and Relationship Strength in Twitter Conversations  
Jin-Yeong Bak, Suin Kim and Alice Oh .............................................. 60

Genre Independent Subgroup Detection in Online Discussion Threads: A Study of Implicit Attitude using Textual Latent Semantics  
Pradeep Dasigi, Weiwei Guo and Mona Diab ...................................... 65

Learning to Temporally Order Medical Events in Clinical Text  
Preethi Raghavan, Albert Lai and Eric Fosler-Lussier .......................... 70

A Context-sensitive, Multi-faceted Model of Lexico-Conceptual Affect  
Tony Veale ...................................................................................... 75
Decoding Running Key Ciphers
   Sravana Reddy and Kevin Knight ...................................................... 80

Using Rejuvenation to Improve Particle Filtering for Bayesian Word Segmentation
   Benjamin Börschinger and Mark Johnson ............................................... 85

Baselines and Bigrams: Simple, Good Sentiment and Topic Classification
   Sida Wang and Christopher Manning ................................................... 90

Automatically Learning Measures of Child Language Development
   Sam Sahakian and Benjamin Snyder ..................................................... 95

A Comparative Study of Target Dependency Structures for Statistical Machine Translation
   Xianchao Wu, Katsuhito Sudoh, Kevin Duh, Hajime Tsukada and Masaaki Nagata .................. 100

Robust Conversion of CCG Derivations to Phrase Structure Trees
   Jonathan K. Kummerfeld, Dan Klein and James R. Curran ......................... 105

Estimating Compact Yet Rich Tree Insertion Grammars
   Elif Yamangil and Stuart Shieber ....................................................... 110

Topic Models for Dynamic Translation Model Adaptation
   Vladimir Eidelman, Jordan Boyd-Graber and Philip Resnik ....................... 115

Detecting Semantic Equivalence and Information Disparity in Cross-lingual Documents
   Yashar Mehdad, Matteo Negri and Marcello Federico ................................ 120

Cross-lingual Parse Disambiguation based on Semantic Correspondence
   Lea Frermann and Francis Bond .......................................................... 125

Learning to Find Translations and Transliterations on the Web
   Joseph Z. Chang, Jason S. Chang and Roger Jyh-Shing Jang ...................... 130

Beefmoves: Dissemination, Diversity, and Dynamics of English Borrowings in a German Hip Hop Forum
   Matt Garley and Julia Hockenmaier ..................................................... 135

Learning the Latent Semantics of a Concept from its Definition
   Weiwei Guo and Mona Diab ................................................................. 140

Unsupervised Semantic Role Induction with Global Role Ordering
   Nikhil Garg and James Henserdon ....................................................... 145

Humor as Circuits in Semantic Networks
   Igor Labutov and Hod Lipson ............................................................ 150

Crowdsourcing Inference-Rule Evaluation
   Naomi Zeichner, Jonathan Berant and Ido Dagan .................................... 156
A Comprehensive Gold Standard for the Enron Organizational Hierarchy
   Apoorv Agarwal, Adinoyi Omuya, Aaron Harnly and Owen Rambow ......................... 161

A Two-step Approach to Sentence Compression of Spoken Utterances
   Dong Wang, Xian Qian and Yang Liu ................................................................. 166

Syntactic Stylometry for Deception Detection
   Song Feng, Ritwik Banerjee and Yejin Choi ....................................................... 171

Transforming Standard Arabic to Colloquial Arabic
   Emad Mohamed, Behrang Mohit and Kemal Oflazer ........................................... 176

Corpus-based Interpretation of Instructions in Virtual Environments
   Luciana Benotti, Martin Villalba, Tessa Lau and Julian Cerruti .......................... 181

Automatically Mining Question Reformulation Patterns from Search Log Data
   Xiaobing Xue, Yu Tao, Daxin Jiang and Hang Li ................................................. 187

Native Language Detection with Tree Substitution Grammars
   Benjamin Swanson and Eugene Charniak ........................................................... 193

Tense and Aspect Error Correction for ESL Learners Using Global Context
   Toshikazu Tajiri, Mamoru Komachi and Yuji Matsumoto .................................... 198

Movie-DiC: a Movie Dialogue Corpus for Research and Development
   Rafael E. Banchs .................................................................................................. 203

Combining Textual Entailment and Argumentation Theory for Supporting Online Debates Interactions
   Elena Cabrio and Serena Villata ........................................................................... 208

Towards the Unsupervised Acquisition of Discourse Relations
   Christian Chiarcos ............................................................................................... 213

Arabic Retrieval Revisited: Morphological Hole Filling
   Kareem Darwish and Ahmed Ali ........................................................................... 218

Extracting and modeling durations for habits and events from Twitter
   Jennifer Williams and Graham Katz ....................................................................... 223

Event Linking: Grounding Event Reference in a News Archive
   Joel Nothman, Matthew Honnibal, Ben Hachey and James R. Curran ..................... 228

Coupling Label Propagation and Constraints for Temporal Fact Extraction
   Yafang Wang, Maximilian Dylla, Marc Spaniol and Gerhard Weikum ...................... 233

Using Search-Logs to Improve Query Tagging
   Kuzman Ganchev, Keith Hall, Ryan McDonald and Slav Petrov .............................. 238

Toward Automatically Assembling Hittite-Language Cuneiform Tablet Fragments into Larger Texts
   Stephen Tyndall .................................................................................................... 243
A Corpus of Textual Revisions in Second Language Writing  
John Lee and Jonathan Webster ................................................................. 248

Coarse Lexical Semantic Annotation with Supersenses: An Arabic Case Study  
Nathan Schneider, Behrang Mohit, Kemal Oflazer and Noah A. Smith .................. 253

Word Epoch Disambiguation: Finding How Words Change Over Time  
Rada Mihalcea and Vivi Nastase ............................................................... 259

Authorship Attribution with Author-aware Topic Models  
Yanir Seroussi, Fabian Bohnert and Ingrid Zukerman .................................. 264

Information-theoretic Multi-view Domain Adaptation  
Pei Yang, Wei Gao, Qi Tan and Kam-Fai Wong .......................................... 270

Efficient Tree-Based Topic Modeling  
Yuening Hu and Jordan Boyd-Graber ......................................................... 275

Learning Better Rule Extraction with Translation Span Alignment  
Jingbo Zhu, Tong Xiao and Chunliang Zhang ............................................. 280

Enhancing Statistical Machine Translation with Character Alignment  
Ning Xi, Guangchao Tang, Xinyu Dai, Shujian Huang and Jiajun Chen ................. 285

Translation Model Size Reduction for Hierarchical Phrase-based Statistical Machine Translation  
Seung-Wook Lee, Dongdong Zhang, Mu Li, Ming Zhou and Hae-Chang Rim .......... 291

Heuristic Cube Pruning in Linear Time  
Andrea Gesmundo, Giorgio Satta and James Henderson ................................ 296

Combining Word-Level and Character-Level Models for Machine Translation Between Closely-Related Languages  
Preslav Nakov and Jörg Tiedemann .......................................................... 301

Improving the IBM Alignment Models Using Variational Bayes  
Darcey Riley and Daniel Gildea ................................................................. 306

Post-ordering by Parsing for Japanese-English Statistical Machine Translation  
Isao Goto, Masao Utiyama and Eiichiro Sumita ......................................... 311

An Exploration of Forest-to-String Translation: Does Translation Help or Hurt Parsing?  
Hui Zhang and David Chiang ...................................................................... 317

Unsupervised Morphology Rivals Supervised Morphology for Arabic MT  
David Stallard, Jacob Devlin, Michael Kayser, Yoong Keok Lee and Regina Barzilay ...... 322

A Meta Learning Approach to Grammatical Error Correction  
Hongsuck Seo, Jonghoon Lee, Seokhwan Kim, Kyusong Lee, Sechun Kang and Gary Geunbae Lee ................................................................. 328
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine Granular Aspect Analysis using Latent Structural Models</td>
<td>Lei Fang and Minlie Huang</td>
<td>333</td>
</tr>
<tr>
<td>Identifying High-Impact Sub-Structures for Convolution Kernels in</td>
<td>Zhaopeng Tu, Yifan He, Jennifer Foster, Josef van Genabith, Qun Liu and Shouxun Lin</td>
<td>338</td>
</tr>
<tr>
<td>Document-level Sentiment Classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploiting Latent Information to Predict Diffusions of Novel Topics</td>
<td>Tsung-Ting Kuo, San-Chuan Hung, Wei-Shih Lin, Nanyun Peng, Shou-De Lin and Wei-Fen Lin</td>
<td>344</td>
</tr>
<tr>
<td>on Social Networks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence Compression with Semantic Role Constraints</td>
<td>Katsumasa Yoshikawa, Ryu Iida, Tsutomu Hirao and Manabu Okumura</td>
<td>349</td>
</tr>
<tr>
<td>Fully Abstractive Approach to Guided Summarization</td>
<td>Pierre-Etienne Genest and Guy Lapalme</td>
<td>354</td>
</tr>
<tr>
<td>Assessing the Effect of Inconsistent Assessors on Summarization</td>
<td>Karolina Owczarzak, Peter A. Rankel, Hoa Trang Dang and John M. Conroy</td>
<td>359</td>
</tr>
<tr>
<td>Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast and Robust Part-of-Speech Tagging Using Dynamic Model Selection</td>
<td>Jinho D. Choi and Martha Palmer</td>
<td>363</td>
</tr>
<tr>
<td>Lemmatisation as a Tagging Task</td>
<td>Andrea Gesmundo and Tanja Samardzic</td>
<td>368</td>
</tr>
<tr>
<td>How Are Spelling Errors Generated and Corrected? A Study of Corrected</td>
<td>Yukino Baba and Hisami Suzuki</td>
<td>373</td>
</tr>
<tr>
<td>and Uncorrected Spelling Errors Using Keystroke Logs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tokenization: Returning to a Long Solved Problem — A Survey,</td>
<td>Rebecca Dridan and Stephan Oepen</td>
<td>378</td>
</tr>
<tr>
<td>Contrastive Experiment, Recommendations, and Toolkit —</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unsupervised Word Segmentation: the Case for Mandarin Chinese</td>
<td>Pierre Magistry and Benoit Sagot</td>
<td>383</td>
</tr>
<tr>
<td>Grammar Error Correction Using Pseudo-Error Sentences and Domain</td>
<td>Kenji Imamura, Kuniko Saito, Kugatsu Sadamitsu and Hitoshi Nishikawa</td>
<td>388</td>
</tr>
<tr>
<td>Adaptation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conference Program

Monday July 9, 2012

(18:00 – 20:30) Poster Session

Wednesday July 11, 2012

Session 7a: (11:00 – 12:30) Parsing

Higher-order Constituent Parsing and Parser Combination
Xiao Chen and Chunyu Kit

Joint Evaluation of Morphological Segmentation and Syntactic Parsing
Reut Tsarfaty, Joakim Nivre and Evelina Andersson

A Comparison of Chinese Parsers for Stanford Dependencies
Wanxiang Che, Valentin Spitkovsky and Ting Liu

A Feature-Rich Constituent Context Model for Grammar Induction
Dave Golland, John DeNero and Jakob Uszkoreit

Session 7b: (11:00 – 12:30) Machine Translation

Private Access to Phrase Tables for Statistical Machine Translation
Nicola Cancedda

Fast and Scalable Decoding with Language Model Look-Ahead for Phrase-based Statistical Machine Translation
Joern Wuebker, Hermann Ney and Richard Zens

Head-Driven Hierarchical Phrase-based Translation
Junhui Li, Zhaopeng Tu, Guodong Zhou and Josef van Genabith

Joint Learning of a Dual SMT System for Paraphrase Generation
Hong Sun and Ming Zhou
Wednesday July 11, 2012 (continued)

Session 7c: (11:00 – 12:30) Relations and events

*A Novel Burst-based Text Representation Model for Scalable Event Detection*
Xin Zhao, Rishan Chen, Kai Fan, Hongfei Yan and Xiaoming Li

*A Graph-based Cross-lingual Projection Approach for Weakly Supervised Relation Extraction*
Seokhwan Kim and Gary Geunbae Lee

*Pattern Learning for Relation Extraction with a Hierarchical Topic Model*
Enrique Alfonseca, Katja Filippova, Jean-Yves Delort and Guillermo Garrido

Session 7d: (11:00 – 12:30) Discourse

*Self-Disclosure and Relationship Strength in Twitter Conversations*
JinYeong Bak, Suin Kim and Alice Oh

*Genre Independent Subgroup Detection in Online Discussion Threads: A Study of Implicit Attitude using Textual Latent Semantics*
Pradeep Dasigi, Weiwei Guo and Mona Diab

*Learning to Temporally Order Medical Events in Clinical Text*
Preethi Raghavan, Albert Lai and Eric Fosler-Lussier

*A Context-sensitive, Multi-faceted Model of Lexico-Conceptual Affect*
Tony Veale

Session 7e: (11:00 – 12:30) Machine Learning

*Decoding Running Key Ciphers*
Sravana Reddy and Kevin Knight

*Using Rejuvenation to Improve Particle Filtering for Bayesian Word Segmentation*
Benjamin Börschinger and Mark Johnson

*Baselines and Bigrams: Simple, Good Sentiment and Topic Classification*
Sida Wang and Christopher Manning
Wednesday July 11, 2012 (continued)

Automatically Learning Measures of Child Language Development
Sam Sahakian and Benjamin Snyder

Session 8a: (16:00 – 17:30) Parsing and Machine Translation

A Comparative Study of Target Dependency Structures for Statistical Machine Translation
Xianchao Wu, Katsuhito Sudoh, Kevin Duh, Hajime Tsukada and Masaaki Nagata

Robust Conversion of CCG Derivations to Phrase Structure Trees
Jonathan K. Kummerfeld, Dan Klein and James R. Curran

Estimating Compact Yet Rich Tree Insertion Grammars
Elif Yamangil and Stuart Shieber

Topic Models for Dynamic Translation Model Adaptation
Vladimir Eidelman, Jordan Boyd-Graber and Philip Resnik

Session 8b: (16:00 – 17:30) Multilinguality

Detecting Semantic Equivalence and Information Disparity in Cross-lingual Documents
Yashar Mehdad, Matteo Negri and Marcello Federico

Cross-lingual Parse Disambiguation based on Semantic Correspondence
Lea Frermann and Francis Bond

Learning to Find Translations and Transliterations on the Web
Joseph Z. Chang, Jason S. Chang and Roger Jyh-Shing Jang

Beefmoves: Dissemination, Diversity, and Dynamics of English Borrowings in a German Hip Hop Forum
Matt Garley and Julia Hockenmaier
Wednesday July 11, 2012 (continued)

Session 8c: (16:00 – 17:30) Meaning

Learning the Latent Semantics of a Concept from its Definition
Weiwei Guo and Mona Diab

Unsupervised Semantic Role Induction with Global Role Ordering
Nikhil Garg and James Henserdon

Humor as Circuits in Semantic Networks
Igor Labutov and Hod Lipson

Crowdsourcing Inference-Rule Evaluation
Naomi Zeichner, Jonathan Berant and Ido Dagan

Session 8d: (16:00 – 17:30) NLP Apps and data

A Comprehensive Gold Standard for the Enron Organizational Hierarchy
Apoorv Agarwal, Adinoyi Omuya, Aaron Harnly and Owen Rambow

A Two-step Approach to Sentence Compression of Spoken Utterances
Dong Wang, Xian Qian and Yang Liu

Syntactic Stylometry for Deception Detection
Song Feng, Ritwik Banerjee and Yejin Choi

Transforming Standard Arabic to Colloquial Arabic
Emad Mohamed, Behrang Mohit and Kemal Oflazer
Wednesday July 11, 2012 (continued)

**Session 8e: (16:00 – 17:30) NLP Apps**

*Corpus-based Interpretation of Instructions in Virtual Environments*
Luciana Benotti, Martin Villalba, Tessa Lau and Julian Cerruti

*Automatically Mining Question Reformulation Patterns from Search Log Data*
Xiaobing Xue, Yu Tao, Daxin Jiang and Hang Li

*Native Language Detection with Tree Substitution Grammars*
Benjamin Swanson and Eugene Charniak

*Tense and Aspect Error Correction for ESL Learners Using Global Context*
Toshikazu Tajiri, Mamoru Komachi and Yuji Matsumoto

Monday July 9, 2012

**(18:00 – 20:30) Poster Session**

*Movie-DiC: a Movie Dialogue Corpus for Research and Development*
Rafael E. Banchs

*Combining Textual Entailment and Argumentation Theory for Supporting Online Debates Interactions*
Elena Cabrio and Serena Villata

*Towards the Unsupervised Acquisition of Discourse Relations*
Christian Chiarcos

*Arabic Retrieval Revisited: Morphological Hole Filling*
Kareem Darwish and Ahmed Ali

*Extracting and modeling durations for habits and events from Twitter*
Jennifer Williams and Graham Katz

*Event Linking: Grounding Event Reference in a News Archive*
Joel Nothman, Matthew Honnibal, Ben Hachey and James R. Curran
Coupling Label Propagation and Constraints for Temporal Fact Extraction
Yafang Wang, Maximilian Dylla, Marc Spaniol and Gerhard Weikum

Using Search-Logs to Improve Query Tagging
Kuzman Ganchev, Keith Hall, Ryan McDonald and Slav Petrov

Toward Automatically Assembling Hittite-Language Cuneiform Tablet Fragments into Larger Texts
Stephen Tyndall

A Corpus of Textual Revisions in Second Language Writing
John Lee and Jonathan Webster

Coarse Lexical Semantic Annotation with Supersenses: An Arabic Case Study
Nathan Schneider, Behrang Mohit, Kemal Oflazer and Noah A. Smith

Word Epoch Disambiguation: Finding How Words Change Over Time
Rada Mihalcea and Vivi Nastase

Authorship Attribution with Author-aware Topic Models
Yanir Seroussi, Fabian Bohnert and Ingrid Zukerman

Information-theoretic Multi-view Domain Adaptation
Pei Yang, Wei Gao, Qi Tan and Kam-Fai Wong

Efficient Tree-Based Topic Modeling
Yuening Hu and Jordan Boyd-Graber

Learning Better Rule Extraction with Translation Span Alignment
Jingbo Zhu, Tong Xiao and Chunliang Zhang

Enhancing Statistical Machine Translation with Character Alignment
Ning Xi, Guangchao Tang, Xinyu Dai, Shujian Huang and Jiajun Chen

Translation Model Size Reduction for Hierarchical Phrase-based Statistical Machine Translation
Seung-Wook Lee, Dongdong Zhang, Mu Li, Ming Zhou and Hae-Chang Rim
Heuristic Cube Pruning in Linear Time
Andrea Gesmundo, Giorgio Satta and James Henderson

Combining Word-Level and Character-Level Models for Machine Translation Between Closely-Related Languages
Preslav Nakov and Jörg Tiedemann

Improving the IBM Alignment Models Using Variational Bayes
Darcey Riley and Daniel Gildea

Post-ordering by Parsing for Japanese-English Statistical Machine Translation
Isao Goto, Masao Utiyama and Eiichiro Sumita

An Exploration of Forest-to-String Translation: Does Translation Help or Hurt Parsing?
Hui Zhang and David Chiang

Unsupervised Morphology Rivals Supervised Morphology for Arabic MT
David Stallard, Jacob Devlin, Michael Kayser, Yoong Keok Lee and Regina Barzilay

A Meta Learning Approach to Grammatical Error Correction
Hongsuck Seo, Jonghoon Lee, Seokhwan Kim, Kyusong Lee, Sechun Kang and Gary Geunbae Lee

Fine Granular Aspect Analysis using Latent Structural Models
Lei Fang and Minlie Huang

Identifying High-Impact Sub-Structures for Convolution Kernels in Document-level Sentiment Classification
Zhaopeng Tu, Yifan He, Jennifer Foster, Josef van Genabith, Qun Liu and Shouxun Lin

Exploiting Latent Information to Predict Diffusions of Novel Topics on Social Networks
Tsung-Ting Kuo, San-Chuan Hung, Wei-Shih Lin, Nanyun Peng, Shou-De Lin and Wei-Fen Lin

Sentence Compression with Semantic Role Constraints
Katsumasa Yoshikawa, Ryu Iida, Tsutomu Hirao and Manabu Okumura

Fully Abstractive Approach to Guided Summarization
Pierre-Etienne Genest and Guy Lapalme
Assessing the Effect of Inconsistent Assessors on Summarization Evaluation  
Karolina Owczarzak, Peter A. Rankel, Hoa Trang Dang and John M. Conroy

Fast and Robust Part-of-Speech Tagging Using Dynamic Model Selection  
Jinho D. Choi and Martha Palmer

Lemmatisation as a Tagging Task  
Andrea Gesmundo and Tanja Samardzic

How Are Spelling Errors Generated and Corrected? A Study of Corrected and Uncorrected Spelling Errors Using Keystroke Logs  
Yukino Baba and Hisami Suzuki

Tokenization: Returning to a Long Solved Problem — A Survey, Contrastive Experiment, Recommendations, and Toolkit —  
Rebecca Dridan and Stephan Oepen

Unsupervised Word Segmentation: the Case for Mandarin Chinese  
Pierre Magistry and Benoît Sagot

Grammar Error Correction Using Pseudo-Error Sentences and Domain Adaptation  
Kenji Imamura, Kuniko Saito, Kugatsu Sadamitsu and Hitoshi Nishikawa