Coling 2010

23rd International Conference on Computational Linguistics

Proceedings of the Conference Volume 2

Chu-Ren Huang and Dan Jurafsky

23 – 27 August 2010
Beijing International Convention Center
Beijing, China

Published by
Tsinghua University Press
All rights reserved for hard copy production.
Block A, Xue Yan Building, Tsinghua University
Beijing, 100084
China

To order hard copies of this proceedings, contact:

Chinese Information Processing Society of China
No.4, Southern Fourth Street
Haidian District, Beijing, 100190
China
Tel: +86-010-62562916
Fax: +86-010-62562916
cips@iscas.ac.cn

Sponsorship

The COLING 2010 very gratefully acknowledges the following commitments in sponsorship:

Platinum Sponsors

-National Natural Science Foundation of China

-Department of Language Information Administration, Ministry of Education, PRC

Gold Sponsor

BaiDu

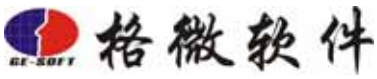
Silver Sponsors

Google

Fujitsu R&D Center CO., LTD.

Microsoft Research

Beijing TRS Information Technology Co., Ltd



Shenyang Globla Envoy software Co.,Ltd.

Supporters

Asian Federation of Natural Language Processing

Institute of Automation
Chinese Academy of Sciences

Institute of Computing Technology
Chinese Academy of Sciences

Institute of Software
Chinese Academy of Sciences

Harbin Institute of Technology

Peking University

Tsinghua University

Preface

You will find in this volume papers from the 23rd International Conference on Computational Linguistics (COLING 2010) held in Beijing, China on August 23-27, 2010 under the auspices of the International Committee on Computational Linguistics (ICCL), and organized by the Chinese Information Processing Society (CIPS) of China. For this prestigious natural language processing conference to be held in China is a significant event for computational linguistics and for colleagues in China, demonstrating both the maturity of our field and the development of academic areas in China.

COLING started as a friendly gathering in New York in 1965, and has grown steadily since. Yet COLINGs aspiration to be a different conference remains the same. COLING strives to maintain its key qualities of embracing different theories and encouraging young scholars in spite of its growing size. A new component introduced at COLING 2010 underlines this quality. A RefreshINGenious (RING) session, organized by Aravind Joshi, our General Chair, allows new and un-orthodox ideas to be presented before they are fully developed in order to generate more discussion and stimulate other new ideas. We hope that this can become an important feature of COLING in the future.

The 155 oral papers included in the hardcopy proceedings published by Tsinghua University Press, as well as the 334 papers included in the electronic proceedings (the same 155 oral papers plus 179 poster papers) are selected from among 815 effective submissions among the more than 840 submissions received. The very selective acceptance rate of 19.02% for oral presentations (155/815 submissions) indicates the extremely high quality of the papers. An additional 21.96% (179/815) are selected for poster presentations to bring the overall acceptance rate to 40.98% (334/815).

We would like to thank the program committee area chairs for their dedicated and efficient review work, and our 738 reviewers for giving us very high quality reviews with a very short turnaround time, allowing us to maintain both the review quality and schedule even given the extraordinary number of submissions. Of course we thank the authors of the 840 papers for submitting their labor of love to COLING. Although we were only able to accept a minority of the submitted papers, we do hope that all authors and reviewers benefit from this process of indirect dialogue. We are especially grateful to the incredibly hard-working team of Stanford volunteers Jenny Finkel, Adam Vogel, and Mengqiu Wang, and HIT volunteers Sam Liang and Lemon Liu, who provided timely and efficient support for the two program chairs at every step of the review and publication processes.

Last but not least, we would like to thank the people who made COLING 2010 and this volume possible. We thank local arrangement committee co-chairs Professor Chengqing Zong and Professor Le Sun for their tireless work which will make COLING-2010 a sure success. Our special appreciation goes to the Chinese Information Processing Society (CIPS) and Professor Youqi Cao for their generous support as the COLING 2010 organizer. Lastly, Professor Qin Lu and Professor Tiejun Zhao should be recognized for their meticulous preparation for editing and publication, which brought this volume to reality.

Chu-Ren Huang and Dan Jurafsky,
COLING 2010 Program Committee Co-chairs

July 8, 2010

COLING 2010 is organized by the Chinese Information Processing Society of China (**CIPS**) and under the auspices of the International Committee on Computational Linguistics (**ICCL**).

General Chair:

Aravind K .Joshi (University of Pennsylvania)

Program Chairs:

Chu-Ren Huang (The Hong Kong Polytechnic University)
Dan Jurafsky (Stanford University)

Advisors to Organizing Committee:

Youqi Cao (The Chinese Information Processing Society of China)
Zhendong Dong (The Chinese Information Processing Society of China)
Changning Huang (Microsoft Research Asia)
Sheng Li (Harbin Institute of Technology)
TianshunYao (Northeastern University)
Shiwen Yu (Peking University)
Zhiwei Feng (Institute of Applied Linguistics, Ministry of Education)
Kaiying Liu (Shanxi University)

OrganizationChairs:

Chengqing Zong (Institute of Automation, Chinese Academy of Sciences)
Le Sun (Institute of Software, Chinese Academy of Sciences)

Publication Chairs:

Qin Lu (The Hong Kong Polytechnic University)
Tiejun Zhao (Harbin Institute of Technology)

Tutorial Chairs:

Dan Gildea (University of Rochester)
Xuanjing Huang (Fudan University)

Workshop Chairs:

Noah Smith (Carnegie Mellon University)
Takenobu Tokunaga (Tokyo Institute of Technology)
Haifeng Wang (BaiDu)

Publicity Chairs:

Hal Daumé III (University of Utah)
Bin Wang (Institute of Computing Technology, Chinese Academy of Sciences)
Minghui Dong (Institute for Infocomm Research)
Monica Monachini (Institute for Computational Linguistics)

Aline Villavicencio (Federal University of Rio Grande do Sul)

Sponsorship Chairs:

Tingting He (Huazhong Normal University)
Hinrich Schütze (University of Stuttgart)
Key-Sun Choi (Korea Advanced Institute of Science and Technology)
Rion Snow (Stanford University)
Takehito Utsuro (University of Tsukuba)
Renata Vieira (Pontifical Catholic University of Rio Grande do Sul)
Hao Yu (Fujitsu(China) R&D Center)

Demo Chairs:

Ting Liu (Harbin Institute of Technology)
Yang Liu (The University of Texas at Dallas)

Program Committee Members/AreaChairs:

Nianwen Xue (Brandeis University)
Rajeev Sangal (India Institute of Information Technology)
Roger Levy (University of California, San Diego)
Justine Cassell (Northwestern University)
Caroline Sporleder (Saarland University)
Gary Geunbae Lee (Pohang University of Science and Technology)
Rosie Jones (Yahoo! Research)
Nicoletta Calzolari (Istituto di Linguistica Computazionale)
Chris Callison-Burch (Johns Hopkins University)
Qun Liu (Institute of Computing Technology)
Pierre Isabelle (National Research Council of Canada)
Sadao Kurohashi (Kyoto University)
Sebastian Padó (Universität Stuttgart)
Takenobu Tokunaga (Tokyo Institute of Technology)
Bo Pang (Yahoo! Research)
Haizhou Li (Institute for Infocomm Research)
David A Smith (University of Massachusetts, Amherst)
Donia Scott (University of Sussex)
Emily Bender (University of Washington)
Ming Zhou (Microsoft Research Asia)

Organization Committee Members:

Juanzi Li (Tsinghua University)
KWONG Olivia (Hong Kong City University)
Bin Sun (Peking University)
Houfeng Wang (Peking University)
Xiaojie Wang (Beijing University of Posts and Telecommunications)
Endong Xun (Beijing Language and Culture University)
Erhong Yang (Beijing Language and Culture University)
Jun Zhao (Institute of Automation, Chinese Academy of Sciences)
Jingbo Zhu (Northeastern University)

Program Committee Members/Reviewers:

Rob Abbott
Takeshi Abekawa
Omri Abend
Steven Abney
Eytan Adar
Guadalupe Aguado
Khalid Al-Kofahi
Cecilia Ovesdotter Alm
Omar Alonso
Bharat Ram Ambati
Massih-Reza Amini
Xiangdong An
Sophia Ananiadou
Jaime Arguello
Masayuki Asahara
Nicholas Asher
Tania Avgustinova
Necip Fazil Ayan
Lakshmi Bai
Jason Baldridge
Timothy Baldwin
Carmen Banea
Srinivas Bangalore
Colin Bannard
Ken Barker
Marco Baroni
Regina Barzilay
John Bateman
Beata Beigman Klebanov
Daisuke Bekki
Nria Bel
Yinon Bentor
Sabine Bergler
Aditya Bhargava
Pushpak Bhattacharyya
Timothy Bickmore
Klinton Bicknell
Alexandra Birch
Philippe Blache
Alan Black
John Blitzer
Michael Bloodgood
Phil Blunsom
Ondrej Bojar
Gemma Boleda
Francis Bond
Johan Bos
Marisa Boston
Pierrette Bouillon
Julien Bourdaillet
Jordan Boyd-Graber
S. R. K. Branavan
Antonio Branco
Eric Breck
Chris Brew
Sabine Buchholz
Stefan Buettcher
Hung Bui
Razvan Bunescu
Aljoscha Burchardt
Stephan Busemann
Miriam Butt
William Byrne
Lynne Cahill
Jamie Callan
Bin Cao
Yunbo Cao
Guiseppe Carenini
Jean Carletta
Marine Carpuat
Xavier Carreras
John Carroll
Ben Carterette
Francisco Casacuberta
Steve Cassidy
Eric Castelli
Alexandru Ceausu
Nick Cercone
Jeong-Won Cha
Vineet Chaitanya
Yllias Chali
Nate Chambers
Baobao Chang
Pi-Chuan Chang
Wanxiang Che
Ciprian Chelba
Aitao Chen
Bin Chen
Boxing Chen
Chien Chin Chen
Harr Chen
Hsin-Hsi Chen
Jinying Chen
Keh-Jiann Chen
Wenliang Chen
Xiao Chen
Charibeth K. Cheng
Colin Cherry
Tee Kiah Chia
David Chiang
Yejin Choi
Kenneth Church
Massimiliano Ciaramita
Philipp Cimiano
Shay Cohen
Kevyn Collins-Thompson
John Conroy
Bonaventura Coppola
Ed Cormany
Dan Cristea
Bruce Croft
Dick Crouch
Montse Cuadros
Silviu-Petru Cucerzan
Oliver Čulo
Aron Culotta
Beatrice Daille
Hercules Dalianis
Sandipan Dandapat
Hoa Dang
Dana Dannells
Dipanjan Das
Sajib Dasgupta
Hal Daume
Dmitry Davidov
Adria de Gispert
Eric De La Clergerie
Valeria de Paiva
Maarten de Rijke
Thierry Declerck
Vera Demberg
Steve DeNeefe
John DeNero
Pascal Denis
Mona Diab
Georgiana Dinu
Pinar Donmez
Iustin Dornescu
Ascander Dost
Antoine Doucet
Doug Downey
Markus Dreyer
Gregory Druck
Xiangyu Duan
Amit Dubey
Kevin Duh
Michael Dukes
Chris Dyer
Marc Dymetman
Markus Egg

Koji Eguchi
David Eichmann
Andreas Eisele
Jacob Eisenstein
Jason Eisner
Noemie Elhadad
Charles Elkan
Micha Elsner
Alexandre entiev
Katrin Erk
Andrea Esuli
Richard Evans
Roger Evans
Patrick Fan
Alex Fang
Benoit Favre
Marcello Federico
Christiane Fellbaum
Katja Filippova
Radu Florian
Sandiway Fong
George Foster
Gil Francopoulo
Stefan Frank
Alex Fraser
Marjorie Freedman
Maria Fuentes
Hagen Fuerstenau
Atsushi Fujii
Sumio Fujita
Pascale Fung
Ryan Gabbard
Evgeniy Gabrilovich
Karthik Gali
Michel Galley
Michael Gamon
Kuzman Ganchev
Jianfeng Gao
Albert Gatt
Eric Gaussier
Ruifang Ge
Byron Georgantopoulos
Pablo Gervas
Sanjukta Ghosh
Daniel Gildea
Dan Gillick
Jesus Gimenez
Kevin Gimpel
Roxana Girju
Sharon Goldwater
Tianxia Gong
Julio Gonzalo
Cyril Goutte
Brigitte Grau
Stephan Greene
Mark Greenwood
Gregory Grefenstette
Ralph Grishman
Cecile Grivaz
Jiafeng Guo
Ben Hachey
Aria Haghighi
Udo Hahn
Jan Hajic
Eva Hajicova
Dilek Hakkani-tur
John Hale
David Hall
Greg Hanneman
Max Harper
Chikara Hashimoto
Ahmed Hassan
Claudia Hauff
Yoshihiko Hayashi
Ben He
daan he
Daqing He
Xiaodong He
Yulan He
John Henderson
Iris Hendrickx
Amac Herdagdelen
Ulf Hermjakob
Raquel Hervas
Dirk Heylen
Andrew Hickl
Graeme Hirst
Jerry Hobbs
Katja Hofmann
Steven Chu Hong Hoi
Kristy Hollingshead
Mark Hopkins
Eric Horvitz
Veronique Hoste
Yunhua Hu
Jimmy Huang
XuanJing Huang
Yifen Huang
Yun Huang
Matt Huenerfauth
Juan Huete
Sarmad Hussain
Rebecca Hwa
Gonzalo Iglesias
Steven Ikier
Diana Inkpen
Kentaro Inui
Elena Irimia
Amy Isard
Abe Ittycheriah
Tatsuya Izuha
Adam Jatowt
Jiwoon Jeon
Heng Ji
Sittichai Jiampojamarn
Daxin Jiang
Hongfei Jiang
Jiepu Jiang
Jing Jiang
Long Jiang
Wenbin Jiang
Valentin Jijkoun
Richard Johansson
Howard Johnson
Rie Johnson
Kristiina Jokinen
Doug Jones
Rosie Jones
Hanmin Jung
Vijay K. Shanker
Min-Yen Kan
Hiroshi Kanayama
Damianos Karakos
Nikiforos Karamanis
Rohit Kate
Tsuneaki Kato
Daisuke Kawahara
Tatsuya Kawahara
Junichi Kazama
Shahram Khadivi
Chloe Kiddon
Jin-Dong Kim
Min Kim
Chunyu Kit
Kevin Knight
Youngjoong Ko
Philipp Koehn
Rob Koeling
Dimitrios Kokkinakis
Greg Kondrak

Moshe Koppel
Valia Kordoni
Lili Kotlerman
Eric Kow
Zornitsa Kozareva
Emiel Kramer
Steven Krauwer
Gerhard Kremer
Canasai Kruengkrai
Lun-Wei Ku
Taku Kudo
Jonas Kuhn
Roland Kuhn
Peter Khnlein
Amba Kulkarni
Ravi Kumar
Shankar Kumar
A Kumaran
Oren Kurland
K. L. Kwok
Olivia Kwong
Wai Lam
Man Lan
Philippe Langlais
Francois Lareau
Martha Larson
Alex Lascarides
Alberto Lavelli
Alon Lavie
Florian Laws
Guy Lebanon
Changki Lee
John Lee
Joo Young Lee
Lillian Lee
Seungwoo Lee
Tan Lee
Jochen Leidner
Yves Lepage
Kevin Lerman
James Lester
Gregor Leusch
Gina-Anne Levow
Hang Li
Jiye Li
Lei Li
Mu Li
Shiqi Li
Shoushan Li
Wei Li
Wen Jie Li
Zhifei Li
Dekang Lin
Lucian Vlad Lita
Diane Litman
Bing Liu
Shui Liu
Ting Liu
Xiaohua Liu
Yan Liu
Yang Liu
Yang Liu
Yupeng Liu
Zhanyi Liu
Zhaopengx Liu
Anna Lobanova
Qiu Long
Adam Lopez
Qin Lu
Yue Lu
Harald Lngen
Xiaoqiang Luo
Tomoya Lwakura
Bin Ma
Qing Ma
Yanjun Ma
Wolfgang Macherey
Nitin Madnani
Hala Maghout
B. Mallikarjun
Gideon Mann
Daniel Marcu
Mitch Marcus
Joseph Mariani
Katja Markert
Erwin Marsi
M. Antonia Marti Antonin
Jean-Claude MARTIN
Andre Martins
Yuval Marton
Yuji Masumoto
Shigeki Matsubara
Tomoko Matsui
Yuji Matsumoto
Takuya Matsuzaki
Evgeny Matusov
Mausam
Arne Mauser
Marshall Mayberry
Diana McCarthy
David McClosky
Ryan McDonald
Tara McIntosh
Kathy McKeown
Susan McRoy
Yashar Mehdad
Qiaozhu Mei
Chris Mellish
Amlia Mendes
Helen Meng
Donald Metzler
Haitao Mi
Jun Miao
Lukas Michelbacher
Jeffrey Micher
Eleni Miltsakaki
David Mimno
Zhaoyan Ming
Shachar Mirkin
Jeff Mitchell
Vibhu Mittal
Yusuke Miyao
Marie-Francine Moens
Dan Moldovan
Diego Moll Aliod
Christof Monz
Raymond J. Mooney
Bob Moore
Roser Morante
Louis-Philippe Morency
Alessandro Moschitti
Isabelle Moulinier
Animesh Mukherjee
Stefan Mller
Art Munson
Dragos Munteanu
Masaki Murata
Vanessa Murdock
Smaranda Muresan
Gabriel Murray
Pradeep Muthukrishnan
Sobha Nair
Testuji Nakagawa
Vivi Nastase
Martina Naughton
Roberto Navigli
Mark-Jan Nederhof
Vasek Nemcik
Ani Nenkova
Hwee Tou Ng

Vincent Ng
Huyen Nguyen Thi Minh
Patrick Nguyen
Alex Niculescu-Mizil
Jian-Yun Nie
Rodney Nielsen
Takashi Ninomiya
Toyoaki Nishida
Cheng Niu
Zheng-Yu Niu
Diarmuid Saghdha
Franz Och
Michael ODonnell
Stephan Oepen
Naoki Okazaki
Lilja Ovrelid
Derya Ozkan
Ulrike Pado
Alexis Palmer
Sinno Jialin Pan
Patrick Pantel
Simone Paolo Ponzetto
Jong Park
Kristen Parton
Marius Pasca
Siddharth Patwardhan
Michael Paul
Soma Paul
Matthias Paulik
Yves Peirsman
Gerald Penn
Marco Pennacchiotti
Bhaskararao Peri
Aina Peris
Wim Peters
Kay Peterson
Slav Petrov
Daniele Pighin
Prasad Pingali
Stelios Piperidis
Guillaume Pitel
Paul Piwek
Luiz Pizzato
Massimo Poesio
Hoifung Poon
Ana-Maria Popescu
Maja Popovic
Christopher Potts
Richard Power
Rashmi Prasad
Laurent Prevot
Matthew Purver
Haoliang Qi
Toh Zhi Qiang
Tao Qin
Yang Qu
Chris Quirk
Bhiksha Raj
S. Rajendran
Dan Ramage
S.V. Ramanan
P. V. S. Rambabu
Owen Rambow
Delip Rao
Ari Rappoport
Paul Rayson
Michaela Regneri
Nils Reiter
Norbert Reithinger
Giuseppe Riccardi
Sebastian Riedel
German Rigau
Hae-Chang Rim
Lucia Rino
Hammam Riza
Horacio Rodrłguez
Gerhard Rolletschek
Lorenza Romano
Mathias Rossignol
Antti-Veikko Rosti
Michael Roth
Alex Rudnick
Marta Ruiz
Kenji Sagae
Horacio Saggion
Harabagiu Sanda
Baskaran Sankaran
Ratna Sanyal
Murat Saraclar
Sudeshna Sarkar
Manabu Sassano
Sekine Satoshi
Roser Sauri
Helmut Schmid
William Schuler
Sabine Schulte im Walde
Ineke Schuurman
Lane Schwartz
Wolfgang Seeker
sekine Sekine
Stephanie Seneff
Violeta Seretan
Hendra Setiawan
Chirag Shah
Dipti Sharma
Dou Shen
Libin Shen
Wade Shen
Kiyoaki Shirai
Eyal Shnarch
Advaith Siddharthan
Candy Sidner
Khalil Simaan
Michel Simard
Kiril Simov
Anil Kumar Singh
Smriti Singh
Samar Sinha
Sharon Small
Jason Smith
Nathaniel Smith
Matthew Snover
Benjamin Snyder
Stephen Soderland
Swapna Somasundaran
Yan Song
Young-In Song
Virach Sornlertlamvanich
Lucia Specia
Jennifer Spenader
Valentin Spitkovsky
Richard Sproat
Ed Stabler
Manfred Stede
Mark Steedman
Amanda Stent
Mark Stevenson
Matthew Stone
Svetlana Stoyanchev
Veselin Stoyanov
Kristina Striegnitz
Michael Strube
Tomek Strzalkowski
Jian Su
Jinsong Su
Keh-Yih Su
Nam Kim Su
Rajen Subba
K. V. Subbarao
L. V. Subramaniam

Zhifang Sui
Eiichiro Sumita
Congkai Sun
Jun Sun
Maosong Sun
Hisami Suzuki
Gyorgy Szarvas
Hiroya Takamura
Koichi Takeuchi
David Talbot
Chew Lim Tan
Jie Tang
Chua Tat-Seng
Paul Tepper
Simone Teufel
Stefan Thater
Mariet Theune
Paul Thomas
Vu Thuy
Joerg Tiedemann
Christoph Tillmann
Ivan Titov
Katrin Tomanek
Yoichi Tomiura
Sara Tonelli
Kentaro Torisawa
Kristina Toutanova
Roy Tromble
Huihsin Tseng
Shu-Chuan Tseng
Benjamin Tsou
Yuta Tsuboi
Jun-ichi Tsujii
Koji Tsukamoto
Yoshimasa Tsuruoka
Dan Tufis
Gokhan Tur
Kiyotaka Uchimoto
Yuya Unno
Lonneke Van der Plas
ielka van der sluis
Josef van Genabith
Gertjan van Noord
Lucy Vanderwende
Sebastian Varges
Sriram Venkatapathy
Marc Verhagen
Yannick Versley
Cristina Vertan
Marta Vila
David Vilar
Aline Villavicencio
Stephan Vogel
Clare Voss
Piek Vossen
Marilyn Walker
Hanna Wallach
Xiaojun Wan
Wenting Wang
Haifeng Wang
Hsin-Min Wang
Kai Wang
Rui Wang
Wei Wang
Xuanhui Wang
Leo Wanner
Nigel Ward
Taro Watanabe
Yotaro Watanabe
Nick Webb
Bonnie Webber
Eric Wehrli
Kilian Weinberger
David Weir
Michael White
Michael Wick
Jan Wiebe
Yorick Wilks
Theresa Wilson
Shuly Wintner
Andreas Witt
Peter Wittek
Magda Wolska
Chung-Hsien Wu
Dan Wu
Dekai Wu
Hua Wu
Xiaoyun Wu
Yunfang Wu
Fei Xia
Rui Xia
Xinyan Xiao
Deyi Xiong
Peng Xu
Ruifeng Xu
Bert Xue
Yongxin Yan
Muyun Yang
Qiang Yang
Ainur Yessenalina
Bei Yu
Jianxing Yu
Kun Yu
Mo Yu
Yisong Yue
Zdenek Zabokrtsky
Taras Zagibalov
Omar Zaidan
Roberto Zanoli
Fabio Zanzotto
Alessandra Zarcone
Richard Zens
Torsten Zesch
Luke Zettlemoyer
Hao Zhang
Hui Zhang
Jiajun Zhang
Min Zhang
Ruiqiang Zhang
Yujie Zhang
Ziqi Zhang
Hai Zhao
Jun Zhao
Bowen Zhou
Guodong Zhou
Yiping Zhou
Conghui Zhu
Jerry Zhu
Tao Zhuang
Heike Zinsmeister
Imed Zitouni
Andreas Zollmann
Chengqing Zong
Ingrid Zukerman

Table of Contents

Testing SDRT's Right Frontier
Stergos Afantenos and Nicholas Asher 1

Identifying Multi-word Expressions by Leveraging Morphological and Syntactic Idiosyncrasy
Hassan Al-Haj and Shuly Wintner 10

Robust Measurement and Comparison of Context Similarity for Finding Translation Pairs
Daniel Andrade, Tetsuya Nasukawa and Junichi Tsujii 19

Multilingual Subjectivity: Are More Languages Better?
Carmen Banea, Rada Mihalcea and Janyce Wiebe 28

Plagiarism Detection across Distant Language Pairs
Alberto Barrón-Cedeño, Paolo Rosso, Eneko Agirre and Gorka Labaka 37

Automatic Detection of Non-deverbal Event Nouns for Quick Lexicon Production
Nuria Bel, Maria Coll and Gabriela Resnik 46

Fast and Accurate Arc Filtering for Dependency Parsing
Shane Bergsma and Colin Cherry 53

A Hierarchical Classifier Applied to Multi-way Sentiment Detection
Adrian Bickerstaffe and Ingrid Zukerman 62

Fluency Constraints for Minimum Bayes-Risk Decoding of Statistical Machine Translation Lattices
Graeme Blackwood, Adrià de Gispert and William Byrne 71

Self-Annotation for fine-grained geospatial relation extraction
Andre Blessing and Hinrich Schütze 80

Top Accuracy and Fast Dependency Parsing is not a Contradiction
Bernd Bohnet 89

Broad Coverage Multilingual Deep Sentence Generation with a Stochastic Multi-Level Realizer
Bernd Bohnet, Leo Wanner, Simon Mill and Alicia Burga 98

Towards an optimal weighting of context words based on distance
Bernard Brosseau-Villeneuve, Jian-Yun Nie and Noriko Kando 107

Measuring the Non-compositionality of Multiword Expressions
Fan Bu, Xiaoyan Zhu and Ming Li 116

A Utility-Driven Approach to Question Ranking in Social QA
Razvan Bunescu and Yunfeng Huang 125

Simultaneous Ranking and Clustering of Sentences: A Reinforcement Approach to Multi-Document Summarization
Xiaoyan Cai, Wenjie Li, You Ouyang and Hong Yan 134

End-to-End Coreference Resolution via Hypergraph Partitioning
Jie Cai and Michael Strube 143

Exploiting Background Knowledge for Relation Extraction
Yee Seng Chan and Dan Roth 152

Jointly Modeling WSD and SRL with Markov Logic
Wanxiang Che and Ting Liu 161

Bipolar Person Name Identification of Topic Documents Using Principal Component Analysis
Chien Chin Chen and Chen-Yuan Wu 170

Emotion Cause Detection with Linguistic Constructions
Ying Chen, Sophia Yat Mei Lee, Shoushan Li and Chu-Ren Huang 179

A Twin-Candidate Based Approach for Event Pronoun Resolution using Composite Kernel
Bin Chen, Jian Su and Chew Lim Tan 188

Unsupervised Synthesis of Multilingual Wikipedia Articles
Yuncong Chen and Pascale Fung 197

Simplicity is Better: Revisiting Single Kernel PPI Extraction
Sung-Pil Choi and Sung-Hyon Myaeng 206

An ontology-driven system for detecting global health events
Nigel Collier, Reiko Matsuda Goodwin, John McCrae, Son Doan, Ai Kawazoe, Mike Conway, Asanee Kawtrakul, Koichi Takeuchi and Dinh Dien 215

Constraining robust constructions for broad-coverage parsing with precision grammars
Bart Cramer and Yi Zhang 223

Local lexical adaptation in Machine Translation through triangulation: SMT helping SMT
Josep Maria Crego, Aurélien Max and François Yvon 232

Automated Translation of Semantic Relationships
Dmitry Davidov and Ari Rappoport 241

Comparison of different algebras for inducing the temporal structure of texts
Pascal Denis and Philippe Muller 250

Generating Learner-Like Morphological Errors in Russian
Markus Dickinson 259

Resolving Object and Attribute Coreference in Opinion Mining
Xiaowen Ding and Bing Liu 268

Entity Disambiguation for Knowledge Base Population
Mark Dredze, Paul McNamee, Delip Rao, Adam Gerber and Tim Finin 277

A Discriminative Latent Variable-Based "DE" Classifier for Chinese-English SMT
Jinhua Du and Andy Way 286

An Empirical Study on Learning to Rank of Tweets
Yajuan Duan, Long Jiang, Tao Qin, Ming Zhou and Heung-Yeung Shum 295

Translation Model Generalization using Probability Averaging for Machine Translation
Nan Duan, Hong Sun and Ming Zhou 304

Mixture Model-based Minimum Bayes Risk Decoding using Multiple Machine Translation Systems
Nan Duan, Mu Li, Dongdong Zhang and Ming Zhou 313

Multi-Sentence Compression: Finding Shortest Paths in Word Graphs
Katja Filippova 322

Enriching Dictionaries with Images from the Internet - Targeting Wikipedia and a Japanese Semantic Lexicon: Lexeed -
Sanae Fujita and Masaaki Nagata 331

Opinosis: A Graph Based Approach to Abstractive Summarization of Highly Redundant Opinions
Kavita Ganesan, ChengXiang Zhai and Jiawei Han 340

EMDC: A Semi-supervised Approach for Word Alignment
Qin Gao, Francisco Guzman and Stephan Vogel 349

A Large Scale Ranker-Based System for Search Query Spelling Correction
Jianfeng Gao, Xiaolong Li, Daniel Micol, Chris Quirk and Xu Sun 358

RTG based surface realisation for TAG
Claire Gardent and Laura Perez-Beltrachini 367

Automatically Learning Source-side Reordering Rules for Large Scale Machine Translation
Dmitriy Genzel 376

Comparing Language Similarity across Genetic and Typologically-Based Groupings
Ryan Georgi, Fei Xia and William Lewis 385

Better Arabic Parsing: Baselines, Evaluations, and Analysis
Spence Green and Christopher D. Manning 394

Paraphrase Alignment for Synonym Evidence Discovery
Gintarè Grigonytè, João Paulo Cordeiro, Gaël Dias, Rumen Moraliyski and Pavel Brazdil . . . 403

Finding the Storyteller: Automatic Spoiler Tagging using Linguistic Cues
Sheng Guo and Naren Ramakrishnan 412

Detection of Simple Plagiarism in Computer Science Papers
Yaakov HaCohen-Kerner, Aharon Tayeb and Natan Ben-Dror 421

A Structured Vector Space Model for Hidden Attribute Meaning in Adjective-Noun Phrases
Matthias Hartung and Anette Frank 430

Hierarchical Phrase-based Machine Translation with Word-based Reordering Model
Katsuhiko Hayashi, Hajime Tsukada, Katsuhito Sudoh, Kevin Duh and Seiichi Yamamoto . . . 439

A Novel Reordering Model Based on Multi-layer Phrase for Statistical Machine Translation
Yanqing He, Yu Zhou, Chengqing Zong and Huilin Wang 447

Standardizing Wordnets in the ISO Standard LMF: Wordnet-LMF for GermaNet
Verena Henrich and Erhard Hinrichs 456

Normal-form parsing for Combinatory Categorial Grammars with generalized composition and type-raising
Julia Hockenmaier and Yonatan Bisk 465

An Empirical Study on Web Mining of Parallel Data
Gumwon Hong, Chi-Ho Li, Ming Zhou and Hae-Chang Rim 474

Enhancing Cross Document Coreference of Web Documents with Context Similarity and Very Large Scale Text Categorization
Jian Huang, Pucktada Treeratpituk, Sarah Taylor and C. Lee Giles 483

Feature-Rich Discriminative Phrase Rescoring for SMT
Fei Huang and Bing Xiang 492

FactRank: Random Walks on a Web of Facts
Alpa Jain and Patrick Pantel 501

Open Entity Extraction from Web Search Query Logs
Alpa Jain and Marco Pennacchiotti 510

Reranking Models in Fine-grained Opinion Analysis
Richard Johansson and Alessandro Moschitti 519

Unsupervised phonemic Chinese word segmentation using Adaptor Grammars
Mark Johnson and Katherine Demuth 528

Data-Driven Parsing with Probabilistic Linear Context-Free Rewriting Systems
Laura Kallmeyer and Wolfgang Maier 537

Learning to Predict Readability using Diverse Linguistic Features
Rohit Kate, Xiaoqiang Luo, Siddharth Patwardhan, Martin Franz, Radu Florian, Raymond Mooney, Salim Roukos and Chris Welty 546

Value for Money: Balancing Annotation Effort, Lexicon Building and Accuracy for Multilingual WSD
Mitesh Khapra, Saurabh Sohoney, Anup Kulkarni and Pushpak Bhattacharyya 555

A Cross-lingual Annotation Projection Approach for Relation Detection
Seokhwan Kim, Minwoo Jeong, Jonghoon Lee and Gary Geunbae Lee 564

Evaluating N-gram based Evaluation Metrics for Automatic Keyphrase Extraction
Su Nam Kim, Timothy Baldwin and Min-Yen Kan 572

Improving the Quality of Text Understanding by Delaying Ambiguity Resolution
Doo Soon Kim, Ken Barker and Bruce Porter 581

Automatic generation of inter-passage links based on semantic similarity
Petr Knoth, Jakub Novotny and Zdenek Zdrahal 590

Dependency-driven Anaphoricity Determination for Coreference Resolution
Fang Kong, Guodong Zhou, Longhua Qian and Qiaoming Zhu 599

Phrase Clustering for Smoothing TM Probabilities - or, How to Extract Paraphrases from Phrase Tables
Roland Kuhn, Boxing Chen, George Foster and Evan Stratford 608

Revisiting Context-based Projection Methods for Term-Translation Spotting in Comparable Corpora
Audrey Laroche and Philippe Langlais 617

Constituent Reordering and Syntax Models for English-to-Japanese Statistical Machine Translation
Young-Suk Lee, Bing Zhao and Xiaoqian Luo 626

Sentiment Classification and Polarity Shifting
Shoushan Li, Sophia Y. M. Lee, Ying Chen, Chu-Ren Huang and Guodong Zhou 635

Improving Corpus Comparability for Bilingual Lexicon Extraction from Comparable Corpora
Bo Li and Eric Gaussier 644

Structure-Aware Review Mining and Summarization
Fangtao Li, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang and Hao Yu . . . 653

Adaptive Development Data Selection for Log-linear Model in Statistical Machine Translation
Mu Li, Yinggong Zhao, Dongdong Zhang and Ming Zhou 662

Learning the Scope of Negation via Shallow Semantic Parsing
Junhui Li, Guodong Zhou, Hongling Wang and Qiaoming Zhu 671

Filtered Ranking for Bootstrapping in Event Extraction
Shasha Liao and Ralph Grishman 680

Exploring variation across biomedical subdomains
Tom Lippincott, Diarmuid Ó Séaghdha, Lin Sun and Anna Korhonen 689

Semantic Role Labeling for News Tweets
Xiaohua Liu, Kuan Li, Bo Han, Ming Zhou, Long Jiang, Zhongyang Xiong and Changning Huang
698

Joint Parsing and Translation
Yang Liu and Qun Liu 707

Semantic Role Features for Machine Translation
Ding Liu and Daniel Gildea 716

TimeML Events Recognition and Classification: Learning CRF Models with Semantic Roles
Hector Llorens, Estela Saquete and Borja Navarro-Colorado 725

Exploiting Structured Ontology to Organize Scattered Online Opinions
Yue Lu, Huizhong Duan, Hongning Wang and ChengXiang Zhai 734

Enhancing Morphological Alignment for Translating Highly Inflected Languages
Minh-Thang Luong and Min-Yen Kan 743

Automatic analysis of semantic similarity in comparable text through syntactic tree matching
Erwin Marsi and Emiel Krahmer 752

Controlling Listening-oriented Dialogue using Partially Observable Markov Decision Processes
Toyomi Meguro, Ryuichiro Higashinaka, Yasuhiro Minami and Kohji Dohsaka 761

Recognising Entailment within Discourse
Shachar Mirkin, Jonathan Berant, Ido Dagan and Eyal Shnarch 770

Evaluating Dependency Representations for Event Extraction
Makoto Miwa, Sampo Pyysalo, Tadayoshi Hara and Jun'ichi Tsujii 779

Entity-Focused Sentence Simplification for Relation Extraction
Makoto Miwa, Rune Sætre, Yusuke Miyao and Jun'ichi Tsujii 788

Using Cross-Lingual Projections to Generate Semantic Role Labeled Annotated Corpus for Urdu - A Resource Poor Language
Smruthi Mukund, Debanjan Ghosh and Rohini Srihari 797

Recognition of Affect, Judgment, and Appreciation in Text
Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka 806

Nonparametric Word Segmentation for Machine Translation
ThuyLinh Nguyen, Stephan Vogel and Noah A. Smith 815

Automatic Discovery of Feature Sets for Dependency Parsing
Peter Nilsson and Pierre Nugues 824

Evaluation of Dependency Parsers on Unbounded Dependencies
Joakim Nivre, Laura Rimell, Ryan McDonald and Carlos Gómez Rodríguez 833

Co-STAR: A Co-training Style Algorithm for Hyponymy Relation Acquisition from Structured and Unstructured Text
Jong-Hoon Oh, Ichiro Yamada, Kentaro Torisawa and Stijn De Saeger 842

Simple and Efficient Algorithm for Approximate Dictionary Matching
Naoaki Okazaki and Jun'ichi Tsujii 851

Latent Mixture of Discriminative Experts for Multimodal Prediction Modeling
Derya Ozkan, Kenji Sagae and Louis-Philippe Morency 860

Text Summarization of Turkish Texts using Latent Semantic Analysis
Makbule Ozsoy, Ilyas Cicekli and Ferda Alpaslan 869

Knowing What to Believe (when you already know something)
Jeff Pasternack and Dan Roth 877

Using Web-scale N-grams to Improve Base NP Parsing Performance
Emily Pitler, Shane Bergsma, Dekang Lin and Kenneth Church 886

Citation Summarization Through Keyphrase Extraction
Vahed Qazvinian, Dragomir R. Radev and Arzucan Ozgur 895

2D Trie for Fast Parsing
Xian Qian, Qi Zhang, Xuangjing Huang and Lide Wu 904

The Bag-of-Opinions Method for Review Rating Prediction from Sparse Text Patterns
Lizhen Qu, Georgiana Ifrim and Gerhard Weikum 913

An Exploration of Features for Recognizing Word Emotion
Changqin Quan and Fuji Ren 922

Inducing Fine-Grained Semantic Classes via Hierarchical and Collective Classification
Altaf Rahman and Vincent Ng 931

Fast, Greedy Model Minimization for Unsupervised Tagging
Sujith Ravi, Ashish Vaswani, Kevin Knight and David Chiang 940

Bringing Active Learning to Life
Ines Rehbein, Josef Ruppenhofer and Alexis Palmer 949

Computing EM-based Alignments of Routes and Route Directions as a Basis for Natural Language Generation
Michael Roth and Anette Frank 958

A Multiple-Domain Ontology Builder
Sara Salem and Samir AbdelRahman 967

Weakly Supervised Morphology Learning for Agglutinating Languages Using Small Training Sets.
Ksenia Shalonova and Bruno Golenia 976

Multi-Document Summarization via the Minimum Dominating Set
Chao Shen and Tao Li 984

Corpus-based Semantic Class Mining: Distributional vs. Pattern-Based Approaches
Shuming Shi, Huibin Zhang, Xiaojie Yuan and Ji-Rong Wen 993

Metaphor Identification Using Verb and Noun Clustering
Ekaterina Shutova, Lin Sun and Anna Korhonen 1002

Explore the Structure of Social Tags by Subsumption Relations
Xiance Si, Zhiyuan Liu and Maosong Sun 1011

Ukwabelana - An open-source morphological Zulu corpus
Sebastian Spiegler, Andrew van der Spuy and Peter A. Flach 1020

EMMA: A novel Evaluation Metric for Morphological Analysis
Sebastian Spiegler and Christian Monson 1029

Modeling Socio-Cultural Phenomena in Discourse
Tomek Strzalkowski, George Aaron Broadwell, Jennifer Stromer-Galley, Samira Shaikh, Sarah Taylor and Nick Webb 1038

Discriminative Induction of Sub-Tree Alignment using Limited Labeled Data
Jun Sun, Min Zhang and Chew Lim Tan 1047

Investigating the cross-linguistic potential of VerbNet-style classification
Lin Sun, Thierry Poibeau, Anna Korhonen and Cedric Messiant 1056

Semi-supervised dependency parsing using generalized tri-training
Anders Søgaard and Christian Rishøj 1065

SemanticRank: Ranking Keywords and Sentences Using Semantic Graphs
George Tsatsaronis, Iraklis Varlamis and Kjetil Nørvåg 1074

Chinese CCGbank: extracting CCG derivations from the Penn Chinese Treebank
Daniel Tse and James R. Curran 1083

Dependency Forest for Statistical Machine Translation
Zhaopeng Tu, Yang Liu, Young-Sook Hwang, Qun Liu and Shouxun Lin 1092

Large Scale Parallel Document Mining for Machine Translation
Jakob Uszkoreit, Jay Ponte, Ashok Popat and Moshe Dubiner 1101

Hungarian Corpus of Light Verb Constructions
Veronika Vincze and János Csirik 1110

Syntax Based Reordering with Automatically Derived Rules for Improved Statistical Machine Translation
Karthik Visweswariah, Jiri Navratil, Jeffrey Sorensen, Vijil Chenthamarakshan and Nandakishore Kambhatla 1119

Efficient Statement Identification for Automatic Market Forecasting
Henning Wachsmuth, Peter Prettenhofer and Benno Stein 1128

Towards a Unified Approach to Simultaneous Single-Document and Multi-Document Summarizations
Xiaojun Wan 1137

"Got You!": Automatic Vandalism Detection in Wikipedia with Web-based Shallow Syntactic-Semantic Modeling
William Yang Wang and Kathleen McKeown 1146

Exploiting Salient Patterns for Question Detection and Question Retrieval in Community-based Question Answering
Kai Wang and Tat-Seng Chua 1155

Probabilistic Tree-Edit Models with Structured Latent Variables for Textual Entailment and Question Answering
Mengqiu Wang and Christopher Manning 1164

A Character-Based Joint Model for Chinese Word Segmentation
Kun Wang, Chengqing Zong and Keh-Yih Su 1173

Near-synonym Lexical Choice in Latent Semantic Space
Tong Wang and Graeme Hirst 1182

Disambiguating Dynamic Sentiment Ambiguous Adjectives
Yunfang Wu and Miaomiao Wen 1191

Joint Tokenization and Translation
Xinyan Xiao, Yang Liu, YoungSook Hwang, Qun Liu and Shouxun Lin 1200

Build Chinese Emotion Lexicons Using A Graph-based Algorithm and Multiple Resources
Ge Xu, Xinfan Meng and Houfeng Wang 1209

A Methodology for Automatic Identification of Nocuous Ambiguity
Hui Yang, Anne De Roeck, Alistair Willis and Bashar Nuseibeh 1218

Contextual Modeling for Meeting Translation Using Unsupervised Word Sense Disambiguation
Mei Yang and Katrin Kirchhoff 1227

A Working Report on Statistically Modeling Dative Variation in Mandarin Chinese
Yao Yao and Feng-hsi Liu 1236

Kernel Slicing: Scalable Online Training with Conjunctive Features
Naoki Yoshinaga and Masaru Kitsuregawa 1245

Discriminative Training for Near-Synonym Substitution
Liang-Chih Yu, Hsiu-Min Shih, Yu-Ling Lai, Jui-Feng Yeh and Chung-Hsien Wu 1254

Estimating Linear Models for Compositional Distributional Semantics
Fabio Massimo Zanzotto, Ioannis Korkontzelos, Francesca Fallucchi and Suresh Manandhar 1263

Grouping Product Features Using Semi-Supervised Learning with Soft-Constraints
Zhongwu Zhai, Bing Liu, Hua Xu and Peifa Jia 1272

Forest-guided Supertagger Training
Yao-zhong Zhang, Takuya Matsuzaki and Jun'ichi Tsujii 1281

Entity Linking Leveraging Automatically Generated Annotation
Wei Zhang, Jian Su, Chew Lim Tan and Wen Ting Wang 1290

A Rhetorical Syntax-Driven Model for Speech Summarization
Jian Zhang and Pascale Fung 1299

Maximum Metric Score Training for Coreference Resolution
Shanheng Zhao and Hwee Tou Ng 1308

Paraphrasing with Search Engine Query Logs
Shiqi Zhao, Haifeng Wang and Ting Liu 1317

Leveraging Multiple MT Engines for Paraphrase Generation
Shiqi Zhao, Haifeng Wang, Xiang Lan and Ting Liu 1326

Resolving Surface Forms to Wikipedia Topics
Yiping Zhou, Lan Nie, Omid Rouhani-Kalleh, Flavian Vasile and Scott Gaffney 1335

Heterogeneous Parsing via Collaborative Decoding
Muhua Zhu, Jingbo Zhu and Tong Xiao 1344

A Monolingual Tree-based Translation Model for Sentence Simplification
Zhemin Zhu, Delphine Bernhard and Iryna Gurevych 1353

A Minimum Error Weighting Combination Strategy for Chinese Semantic Role Labeling
Tao Zhuang and Chengqing Zong 1362

Detecting Speech Repairs Incrementally Using a Noisy Channel Approach
Simon Zwarts, Mark Johnson and Robert Dale 1371

Syntactic Scope Resolution in Uncertainty Analysis
Lilja Øvrelid, Erik Velldal and Stephan Oepen 1379