Table of Contents

Preface: General Chair xiii

Preface: Program Committee Co-Chairs xv

Organizers xvii

Program Committee xxi

Conference Program xxv

Combination of Arabic Preprocessing Schemes for Statistical Machine Translation
Fatiha Sadat and Nizar Habash 1

Going Beyond AER: An Extensive Analysis of Word Alignments and Their Impact on MT
Necip Fazil Ayan and Bonnie J. Dorr 9

Unsupervised Topic Modelling for Multi-Party Spoken Discourse
Matthew Purver, Konrad P. Körding, Thomas L. Griffiths and Joshua B. Tenenbaum 17

Minimum Cut Model for Spoken Lecture Segmentation
Igor Malioutov and Regina Barzilay 25

Bootstrapping Path-Based Pronoun Resolution
Shane Bergsma and Dekang Lin 33

Kernel-Based Pronoun Resolution with Structured Syntactic Knowledge
Xiaofeng Yang, Jian Su and Chew Lim Tan 41

A Finite-State Model of Human Sentence Processing
Jihyun Park and Chris Brew 49

Acceptability Prediction by Means of Grammaticality Quantification
Philippe Blache, Barbara Hemforth and Stéphane Rauzy 57

Discriminative Word Alignment with Conditional Random Fields
Phil Blunsom and Trevor Cohn 65

Named Entity Transliteration with Comparable Corpora
Richard Sproat, Tao Tao and ChengXiang Zhai 73

Extracting Parallel Sub-Sentential Fragments from Non-Parallel Corpora
Dragos Stefan Munteanu and Daniel Marcu 81

Estimating Class Priors in Domain Adaptation for Word Sense Disambiguation
Yee Seng Chan and Hwee Tou Ng 89

Ensemble Methods for Unsupervised WSD
Samuel Brody, Roberto Navigli and Mirella Lapata 97

Meaningful Clustering of Senses Helps Boost Word Sense Disambiguation Performance
Roberto Navigli 105

Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations
Patrick Pantel and Marco Pennacchiotti 113

Modeling Commonality among Related Classes in Relation Extraction
GuoDong Zhou, Jian Su and Min Zhang 121

Relation Extraction Using Label Propagation Based Semi-Supervised Learning
Jinxiu Chen, Donghong Ji, Chew Lim Tan and Zhengyu Niu 129

Polarized Unification Grammars
Sylvain Kahane 137

Partially Specified Signatures: A Vehicle for Grammar Modularity
Yael Cohen-Sygal and Shuly Wintner 145

Morphology-Syntax Interface for Turkish LFG
Özlem Çetinoğlu and Kemal Oflazer 153

PCFGs with Syntactic and Prosodic Indicators of Speech Repairs
John Hale, Izhak Shafran, Lisa Yung, Bonnie Dorr, Mary Harper, Anna Krasnyanskaya, Matthew Lease, Yang Liu, Brian Roark, Matthew Snover and Robin Stewart 161

Dependency Parsing of Japanese Spoken Monologue Based on Clause Boundaries
Tomohiro Ohno, Shigeki Matsubara, Hideki Kashioka, Takehiko Maruyama and Yasuyoshi Inagaki 169

Trace Prediction and Recovery with Unlexicalized PCFGs and Slash Features
Helmut Schmid 177

Learning More Effective Dialogue Strategies Using Limited Dialogue Move Features
Matthew Frampton and Oliver Lemon 185

Dependencies between Student State and Speech Recognition Problems in Spoken Tutoring Dialogues
Mihai Rotaru and Diane J. Litman 193

Learning the Structure of Task-Driven Human-Human Dialogs
Srinivas Bangalore, Giuseppe Di Fabbrizio and Amanda Stent 201

Semi-Supervised Conditional Random Fields for Improved Sequence Segmentation and Labeling
Feng Jiao, Shaojun Wang, Chi-Hoon Lee, Russell Greiner and Dale Schuurmans 209

Training Conditional Random Fields with Multivariate Evaluation Measures
Jun Suzuki, Erik McDermott and Hideki Isozaki 217

Approximation Lasso Methods for Language Modeling
Jianfeng Gao, Hisami Suzuki and Bin Yu 225

Automated Japanese Essay Scoring System based on Articles Written by Experts
Tsunenori Ishioka and Masayuki Kameda 233

A Feedback-Augmented Method for Detecting Errors in the Writing of Learners of English
Ryo Nagata, Atsuo Kawai, Koichiro Morihiro and Naoki Isu 241

Correcting ESL Errors Using Phrasal SMT Techniques
Chris Brockett, William B. Dolan and Michael Gamon 249

Graph Transformations in Data-Driven Dependency Parsing
Jens Nilsson, Joakim Nivre and Johan Hall 257

Learning to Generate Naturalistic Utterances Using Reviews in Spoken Dialogue Systems
Ryuichiro Higashinaka, Rashmi Prasad and Marilyn A. Walker 265

Measuring Language Divergence by Intra-Lexical Comparison
T. Mark Ellison and Simon Kirby 273

Enhancing Electronic Dictionaries with an Index Based on Associations
Olivier Ferret and Michael Zock 281

Guiding a Constraint Dependency Parser with Supertags
Kilian Foth, Tomas By and Wolfgang Menzel 289

Efficient Unsupervised Discovery of Word Categories Using Symmetric Patterns and High Frequency Words
Dmitry Davidov and Ari Rappoport 297

Bayesian Query-Focused Summarization
Hal Daumé III and Daniel Marcu 305

Expressing Implicit Semantic Relations without Supervision
Peter D. Turney 313

Hybrid Parsing: Using Probabilistic Models as Predictors for a Symbolic Parser
Kilian A. Foth and Wolfgang Menzel 321

Error Mining in Parsing Results
Benoît Sagot and Éric de La Clergerie 329

Reranking and Self-Training for Parser Adaptation
David McClosky, Eugene Charniak and Mark Johnson 337

Automatic Classification of Verbs in Biomedical Texts
Anna Korhonen, Yuval Krymolowski and Nigel Collier 345

Selection of Effective Contextual Information for Automatic Synonym Acquisition
Masato Hagiwara, Yasuhiro Ogawa and Katsuhiko Toyama 353

Scaling Distributional Similarity to Large Corpora
James Gorman and James R. Curran 361

Extractive Summarization using Inter- and Intra- Event Relevance
Wenjie Li, Mingli Wu, Qin Lu, Wei Xu and Chunfa Yuan 369

Models for Sentence Compression: A Comparison across Domains, Training Requirements and Evaluation Measures
James Clarke and Mirella Lapata 377

A Bottom-Up Approach to Sentence Ordering for Multi-Document Summarization
Danushka Bollegala, Naoaki Okazaki and Mitsuru Ishizuka 385

Learning Event Durations from Event Descriptions
Feng Pan, Rutu Mulkar and Jerry R. Hobbs 393

Automatic Learning of Textual Entailments with Cross-Pair Similarities
Fabio Massimo Zanzotto and Alessandro Moschitti 401

An Improved Redundancy Elimination Algorithm for Underspecified Representations
Alexander Koller and Stefan Thater 409

Integrating Syntactic Priming into an Incremental Probabilistic Parser, with an Application to Psycholinguistic Modeling
Amit Dubey, Frank Keller and Patrick Sturt 417

A Fast, Accurate Deterministic Parser for Chinese
Mengqiu Wang, Kenji Sagae and Teruko Mitamura 425

Learning Accurate, Compact, and Interpretable Tree Annotation
Slav Petrov, Leon Barrett, Romain Thibaux and Dan Klein 433

Semi-Supervised Learning of Partial Cognates Using Bilingual Bootstrapping
Oana Frunza and Diana Inkpen 441

Direct Word Sense Matching for Lexical Substitution
Ido Dagan, Oren Glickman, Alfio Gliozzo, Efrat Marmorshtein and Carlo Strapparava 449

An Equivalent Pseudoword Solution to Chinese Word Sense Disambiguation
Zhimao Lu, Haifeng Wang, Jianmin Yao, Ting Liu and Sheng Li 457

Improving the Scalability of Semi-Markov Conditional Random Fields for Named Entity Recognition
Daisuke Okanohara, Yusuke Miyao, Yoshimasa Tsuruoka and Jun'ichi Tsujii 465

Factorizing Complex Models: A Case Study in Mention Detection
Radu Florian, Hongyan Jing, Nanda Kambhatla and Imed Zitouni 473

Segment-Based Hidden Markov Models for Information Extraction
Zhenmei Gu and Nick Cercone 481

A DOM Tree Alignment Model for Mining Parallel Data from the Web
Lei Shi, Cheng Niu, Ming Zhou and Jianfeng Gao 489

QuestionBank: Creating a Corpus of Parse-Annotated Questions
John Judge, Aoife Cahill and Josef van Genabith 497

Creating a CCGbank and a Wide-Coverage CCG Lexicon for German
Julia Hockenmaier 505

Improved Discriminative Bilingual Word Alignment
Robert C. Moore, Wen-tau Yih and Andreas Bode 513

Maximum Entropy Based Phrase Reordering Model for Statistical Machine Translation
Deyi Xiong, Qun Liu and Shouxun Lin 521

Distortion Models for Statistical Machine Translation
Yaser Al-Onaizan and Kishore Papineni 529

A Study on Automatically Extracted Keywords in Text Categorization
Anette Hulth and Beáta B. Megyesi 537

A Comparison and Semi-Quantitative Analysis of Words and Character-Bigrams as Features in Chinese Text Categorization
Jingyang Li, Maosong Sun and Xian Zhang 545

Exploiting Comparable Corpora and Bilingual Dictionaries for Cross-Language Text Categorization
Alfio Gliozzo and Carlo Strapparava 553

A Progressive Feature Selection Algorithm for Ultra Large Feature Spaces
Qi Zhang, Fuliang Weng and Zhe Feng 561

Annealing Structural Bias in Multilingual Weighted Grammar Induction
Noah A. Smith and Jason Eisner 569

Maximum Entropy Based Restoration of Arabic Diacritics
Imed Zitouni, Jeffrey S. Sorensen and Ruhi Sarikaya 577

An Iterative Implicit Feedback Approach to Personalized Search
Yuanhua Lv, Le Sun, Junlin Zhang, Jian-Yun Nie, Wan Chen and Wei Zhang 585

The Effect of Translation Quality in MT-Based Cross-Language Information Retrieval
Jiang Zhu and Haifeng Wang 593

A Comparison of Document, Sentence, and Term Event Spaces
Catherine Blake 601

Tree-to-String Alignment Template for Statistical Machine Translation
Yang Liu, Qun Liu and Shouxun Lin 609

Incorporating Speech Recognition Confidence into Discriminative Named Entity Recognition of Speech Data
Katsuhito Sudoh, Hajime Tsukada and Hideki Isozaki 617

Exploiting Syntactic Patterns as Clues in Zero-Anaphora Resolution
Ryu Iida, Kentaro Inui and Yuji Matsumoto 625

Self-Organizing n-gram Model for Automatic Word Spacing
Seong-Bae Park, Yoon-Shik Tae and Se-Young Park 633

Concept Unification of Terms in Different Languages for IR
Qing Li, Sung-Hyon Myaeng, Yun Jin and Bo-yeong Kang 641

Word Alignment in English-Hindi Parallel Corpus Using Recency-Vector Approach: Some Studies
Niladri Chatterjee and Saumya Agrawal 649

Extracting Loanwords from Mongolian Corpora and Producing a Japanese-Mongolian Bilingual Dictionary
Badam-Osor Khaltar, Atsushi Fujii and Tetsuya Ishikawa 657

An Unsupervised Morpheme-Based HMM for Hebrew Morphological Disambiguation
Meni Adler and Michael Elhadad 665

Contextual Dependencies in Unsupervised Word Segmentation
Sharon Goldwater, Thomas L. Griffiths and Mark Johnson 673

MAGEAD: A Morphological Analyzer and Generator for the Arabic Dialects
Nizar Habash and Owen Rambow 681

Noun Phrase Chunking in Hebrew: Influence of Lexical and Morphological Features
Yoav Goldberg, Meni Adler and Michael Elhadad 689

Multi-Tagging for Lexicalized-Grammar Parsing
James R. Curran, Stephen Clark and David Vadas 697

Guessing Parts-of-Speech of Unknown Words Using Global Information
Tetsuji Nakagawa and Yuji Matsumoto 705

A Clustered Global Phrase Reordering Model for Statistical Machine Translation
Masaaki Nagata, Kuniko Saito, Kazuhide Yamamoto and Kazuteru Ohashi 713

A Discriminative Global Training Algorithm for Statistical MT
Christoph Tillmann and Tong Zhang 721

Phoneme-to-Text Transcription System with an Infinite Vocabulary
Shinsuke Mori, Daisuke Takuma and Gakuto Kurata 729

Automatic Generation of Domain Models for Call-Centers from Noisy Transcriptions
Shourya Roy and L Venkata Subramaniam 737

Proximity in Context: An Empirically Grounded Computational Model of Proximity for Processing Topological Spatial Expressions
John D. Kelleher, Geert-Jan M. Kruijff and Fintan J. Costello 745

Machine Learning of Temporal Relations
Inderjeet Mani, Marc Verhagen, Ben Wellner, Chong Min Lee and James Pustejovsky 753

An End-to-End Discriminative Approach to Machine Translation
Percy Liang, Alexandre Bouchard-Côté, Dan Klein and Ben Taskar 761

Semi-Supervised Training for Statistical Word Alignment
Alexander Fraser and Daniel Marcu 769

Left-to-Right Target Generation for Hierarchical Phrase-Based Translation
Taro Watanabe, Hajime Tsukada and Hideki Isozaki 777

You Can't Beat Frequency (Unless You Use Linguistic Knowledge) – A Qualitative Evaluation of Association Measures for Collocation and Term Extraction
Joachim Wermter and Udo Hahn 785

Ontologizing Semantic Relations
Marco Pennacchiotti and Patrick Pantel 793

Semantic Taxonomy Induction from Heterogenous Evidence
Rion Snow, Daniel Jurafsky and Andrew Y. Ng 801

Names and Similarities on the Web: Fact Extraction in the Fast Lane
Marius Paşca, Dekang Lin, Jeffrey Bigham, Andrei Lifchits and Alpa Jain 809

Weakly Supervised Named Entity Transliteration and Discovery from Multilingual Comparable Corpora
Alexandre Klementiev and Dan Roth 817

A Composite Kernel to Extract Relations between Entities with Both Flat and Structured Features
Min Zhang, Jie Zhang, Jian Su and GuoDong Zhou 825

Japanese Dependency Parsing Using Co-Occurrence Information and a Combination of Case Elements
Takeshi Abekawa and Manabu Okumura 833

Answer Extraction, Semantic Clustering, and Extractive Summarization for Clinical Question Answering
Dina Demner-Fushman and Jimmy Lin 841

Discovering Asymmetric Entailment Relations between Verbs Using Selectional Preferences
Fabio Massimo Zanzotto, Marco Pennacchiotti and Maria Teresa Pazienza 849

Event Extraction in a Plot Advice Agent
Harry Halpin and Johanna D. Moore 857

An All-Subtrees Approach to Unsupervised Parsing
Rens Bod 865

Advances in Discriminative Parsing
Joseph Turian and I. Dan Melamed 873

Prototype-Driven Grammar Induction
Aria Haghighi and Dan Klein 881

Exploring Correlation of Dependency Relation Paths for Answer Extraction
Dan Shen and Dietrich Klakow 889

Question Answering with Lexical Chains Propagating Verb Arguments
Adrian Novischi and Dan Moldovan 897

Methods for Using Textual Entailment in Open-Domain Question Answering
Sanda Harabagiu and Andrew Hickl 905

Using String-Kernels for Learning Semantic Parsers
Rohit J. Kate and Raymond J. Mooney 913

A Bootstrapping Approach to Unsupervised Detection of Cue Phrase Variants
Rashid M. Abdalla and Simone Teufel 921

Semantic Role Labeling via FrameNet, VerbNet and PropBank
Ana-Maria Giuglea and Alessandro Moschitti 929

Multilingual Legal Terminology on the Jibiki Platform: The LexALP Project
Gilles Sérasset, Francis Brunet-Manquat and Elena Chiocchetti 937

Leveraging Reusability: Cost-Effective Lexical Acquisition for Large-Scale Ontology Translation
G. Craig Murray, Bonnie J. Dorr, Jimmy Lin, Jan Hajič and Pavel Pecina 945

Accurate Collocation Extraction Using a Multilingual Parser
Violeta Seretan and Eric Wehrli 953

Scalable Inference and Training of Context-Rich Syntactic Translation Models
Michel Galley, Jonathan Graehl, Kevin Knight, Daniel Marcu, Steve DeNeefe, Wei Wang and Ignacio Thayer 961

Modelling Lexical Redundancy for Machine Translation
David Talbot and Miles Osborne 969

Empirical Lower Bounds on the Complexity of Translational Equivalence
Benjamin Wellington, Sonjia Waxmonsky and I. Dan Melamed 977

A Hierarchical Bayesian Language Model Based On Pitman-Yor Processes
Yee Whye Teh 985

A Phonetic-Based Approach to Chinese Chat Text Normalization
Yunqing Xia, Kam-Fai Wong and Wenjie Li 993

Discriminative Pruning of Language Models for Chinese Word Segmentation
Jianfeng Li, Haifeng Wang, Dengjun Ren and Guohua Li 1001

Novel Association Measures Using Web Search with Double Checking
Hsin-Hsi Chen, Ming-Shun Lin and Yu-Chuan Wei 1009

Semantic Retrieval for the Accurate Identification of Relational Concepts in Massive Textbases
Yusuke Miyao, Tomoko Ohta, Katsuya Masuda, Yoshimasa Tsuruoka, Kazuhiro Yoshida, Takashi Ninomiya and Jun'ichi Tsujii 1017

Exploring Distributional Similarity Based Models for Query Spelling Correction
Mu Li, Muhua Zhu, Yang Zhang and Ming Zhou 1025

Robust PCFG-Based Generation Using Automatically Acquired LFG Approximations
Aoife Cahill and Josef van Genabith 1033

Incremental Generation of Spatial Referring Expressions in Situated Dialog
John D. Kelleher and Geert-Jan M. Kruijff 1041

Learning to Predict Case Markers in Japanese
Hisami Suzuki and Kristina Toutanova 1049

Are These Documents Written from Different Perspectives? A Test of Different Perspectives Based on Statistical Distribution Divergence
Wei-Hao Lin and Alexander Hauptmann 1057

Word Sense and Subjectivity
Janyce Wiebe and Rada Mihalcea 1065

Improving QA Accuracy by Question Inversion
John Prager, Pablo Duboue and Jennifer Chu-Carroll 1073

Reranking Answers for Definitional QA Using Language Modeling
Yi Chen, Ming Zhou and Shilong Wang 1081

Highly Constrained Unification Grammars
Daniel Feinstein and Shuly Wintner 1089

A Polynomial Parsing Algorithm for the Topological Model: Synchronizing Constituent and Dependency Grammars, Illustrated by German Word Order Phenomena
Kim Gerdes and Sylvain Kahane 1097

Stochastic Language Generation Using WIDL-Expressions and its Application in Machine Translation and Summarization
Radu Soricut and Daniel Marcu 1105

Learning to Say It Well: Reranking Realizations by Predicted Synthesis Quality
Crystal Nakatsu and Michael White 1113

An Effective Two-Stage Model for Exploiting Non-Local Dependencies in Named Entity Recognition
Vijay Krishnan and Christopher D. Manning 1121

Learning Transliteration Lexicons from the Web
Jin-Shea Kuo, Haizhou Li and Ying-Kuei Yang 1129

Punjabi Machine Transliteration
M.G. Abbas Malik 1137

Multilingual Document Clustering: An Heuristic Approach Based on Cognate Named Entities
Soto Montalvo, Raquel Martínez, Arantza Casillas and Víctor Fresno 1145

Time Period Identification of Events in Text
Taichi Noro, Takashi Inui, Hiroya Takamura and Manabu Okumura 1153

Optimal Constituent Alignment with Edge Covers for Semantic Projection
Sebastian Padó and Mirella Lapata 1161

Utilizing Co-Occurrence of Answers in Question Answering
Min Wu and Tomek Strzalkowski 1169

Author Index 1177

Preface: General Chair

I am honoured to write the first few words of these Proceedings, as General Chair of COLING/ACL 2006 in Sydney, Australia. As we know, this is just the third time in their history that the two traditionally major events in Computational Linguistics, COLING and ACL – organised respectively by ICCL (International Committee on Computational Linguistics) and ACL (the Association for Computational Linguistics) – are joined in one combined conference, after Stanford in 1984 and Montreal in 1998. I was lucky to attend both those wonderful events and would have never imagined to be "in charge" of the next one, the first of the new millennium!

When I accepted, I knew I didn't have real work to do in this position, apart from mediate – if necessary – among the several "real workers", the various Chairs. I must say now that my work was even easier than foreseen, because of the wonderful teamwork of all the COLING/ACL group.

In this joint Conference we have tried to maintain the spirit of both COLING and ACL, but the combination will inevitably have its own personality, in a mixture that is more than the simple sum of the two. Part of its character will be due to the location, for the first time – for both conferences – in Australia. For this reason we decided to have a member of AFNLP (the Asian Federation of Natural Language Processing) on the Advisory Board and to give particular attention and visibility to the Asia-Pacific context, communities and languages. We sincerely thank both the AFNLP-Nagao Fund for providing financial support for those presenting Asian NLP research, and ALTA (the Australasian Language Technology Association) for their local support.

It is my task here – but I should say my pleasure – to express gratitude to all those without whom this conference would not exist, and I think I can do that on behalf of all participants.

My biggest thanks go to all the Chairs, for their invaluable effort and dedication which made this Conference possible.

First of all the two Program Chairs: Claire Cardie and Pierre Isabelle, who did a tremendous job, managing so many submissions and taking care of both regular papers and posters, and the two Local Arrangements Chairs: Robert Dale and Cécile Paris, who have succeeded in keeping so many details under control, in such a smooth way as if everything were natural and effortless for them.

And all the others, for their precious, competent and hard work: the Workshops Chair: Suzanne Stevenson; the Student Workshop Chair: Rebecca Hwa; the Tutorials Chair: Claire Gardent; the Interactive Presentations Chair: James Curran; the Publications Chair: Olivia Kwong; the two Sponsorship Chairs: Steven Krauwer (International) and Dominique Estival (Australia); the Mentoring Chair: Richard Power, who kindly accepted to do this for the second time; the Publicity Chair: Tim Baldwin; the Exhibits Chair: Menno van Zaanen; the Student Volunteers coordinator: Priscilla Rasmussen, giving often advice to all of us as ACL business manager; the webmasters: Andrew Lampert and Brett Powley; and finally Judy Potter and her team from Well Done Events for managing registrations and assisting in the local organisation.

I warmly thank the Advisory Board – composed of four ICCL, four ACL, and one AFNLP members – to whom we resorted for suggestions on important and sometimes delicate issues: Sandra Carberry, Eva Hajicova, Aravind Joshi, Martin Kay, Kathleen McCoy, Martha Palmer, Priscilla Rasmussen, Benjamin T'sou, Jun'ichi Tsujii.

I express my gratitude to all the sponsors for their great support to the conference.

I thank all the organizers of the so numerous surrounding workshops, tutorials, and other co-located events – conferences, workshops, summer school – adding value to the main conference, creating

altogether probably the biggest ever happening in Computational Linguistics.

My thanks to the area chairs, the reviewers, the invited speakers, the authors of the various presentations, in particular the students who enter with enthusiasm in such an exciting field, all the participants who will often make a long trip to be present at COLING/ACL 2006, and all those who contributed in many ways to a success of the conference.

And I finally thank both ICCL and ACL for having decided to join forces again in such a great enterprise. COLING/ACL 2006 will be, I'm sure, an exciting, stimulating and inspiring event for all of you.

Enjoy COLING/ACL 2006! . . . and consider that some of the youngest here do not know it yet, but they will be chairing the next joint events in a few years.

Nicoletta Calzolari
COLING/ACL 2006 General Chair
June 2006

Preface: Program Committee Co-Chairs

This conference represents just the third time in their 40+ year history that the two premier conferences in natural language processing, computational linguistics, and language technology have merged for a joint COLING/ACL event; and it's the first time that the joint conference will be held in the southern hemisphere. It is fitting then, that we received a record number of 630 submissions from 40+ countries: 39% from 13 countries in Asia, 29% from 17 countries in Europe, 25% from Canada and the United States, 4% from Australia and New Zealand, 2% from 4 countries in the Middle East, and less than 1% from South America (Brazil) and Africa (South Africa and Tunisia). Of the 630 submissions, 23% were accepted for paper presentations and an additional 20% for poster presentations.

Our rough estimate of the amount of work that went just into preparing the submissions, final versions and on-site presentations for this year's main program exceeds 32 person/years.[1] If we include workshops and EMNLP, this figure is probably doubled. Thanks to everyone who submitted their research to the conference!

Much of the work in putting together the main program of papers and posters was done, of course, by our tireless area chairs and reviewers (of which we had 19 and 384, respectively). A tribute to their joint efforts is the fact that we obtained a 100% response rate for reviews – over 100% actually, since a few crazy souls offered unsolicited or extra reviews.

COLING/ACL 2006 spans five days with the traditional COLING "excursion day" on day three. The remaining four days of the conference include plenary sessions, four parallel paper sessions, the student research workshop, and two evening poster sessions. The ACL Lifetime Achievement Award will also be bestowed on its fifth recipient in a plenary session, followed by an invited talk by the esteemed award winner. A Best Paper Award will be announced in a plenary session at the end of the conference. We would like to especially thank our two invited speakers, Daniel Marcu and Sally McConnell-Ginet.

In honor of the joint conference's location, we have planned a special Asian language event for Thursday morning that consists of paper presentations of the top four Asian language papers followed by a plenary panel focusing on issues in Asian language processing, and ending with the presentation of the Best Asian Language Paper Award. We offer special thanks to our three distinguished panelists – Pushpak Bhattacharyya, Benjamin T'sou, and Jun'ichi Tsujii – and to Aravind Joshi, who expertly organized the panel.

Finally, we thank the ACL and ICCL conference oversight committee, for advice of all sorts along the way; and Rich Gerber, the START conference system developer, who answered our countless questions at all hours of the day and night.

After all of this work, by so many people, we are very much looking forward to sitting back and enjoying the conference with you in Sydney in July!

Claire Cardie
Pierre Isabelle
June 2006

[1] We assume an average of 8 days of work to prepare each one of about 630 submissions to the COLING-ACL 2006 main program; and an average of 5 days of work to produce final versions for each one of the 267 accepted contributions.

Organizers

General Chair:

Nicoletta Calzolari, Istituto di Linguistica Computazionale – CNR, Italy

Program Committee Co-Chairs:

Claire Cardie, Cornell University, USA
Pierre Isabelle, National Research Council of Canada, Canada

Tutorials Chair:

Claire Gardent, CNRS/LORIA, France

Workshops Chair:

Suzanne Stevenson, University of Toronto, Canada

Workshops Program Committee:

Ann Copestake, University of Cambridge, UK
Pascale Fung, Hong Kong University of Science and Technology, Hong Kong
Jamie Henderson, University of Edinburgh, UK
Ingrid Zukerman, Monash University, Australia

Interactive Presentations Chair:

James Curran, University of Sydney, Australia

Publications Chair:

Olivia Kwong, City University of Hong Kong, Hong Kong

Sponsorship Chairs:

Steven Krauwer, ELSNET / UiL OTS, The Netherlands
Dominique Estival, Appen Pty Limited, Australia

Exhibits Chair:

Menno van Zaanen, Macquarie University, Australia

Mentoring Service Chair:

Richard Power, The Open University, UK

Publicity Chair:

Timothy Baldwin, University of Melbourne, Australia

Student Research Workshop:

Rebecca Hwa, University of Pittsburgh, USA
Marine Carpuat, Hong Kong University of Science and Technology, Hong Kong
Kevin Duh, University of Washington, USA

Local Organization Chairs:

Robert Dale, Macquarie University, Australia
Cécile Paris, CSIRO ICT Centre, Australia

Local Organizing Advisory Committee:

John Debenham, University of Technology, Sydney, Australia
Jon Patrick, University of Sydney, Australia
Raymond Wong, University of New South Wales, Australia

Student Volunteers Coordinator:

Priscilla Rasmussen, Association for Computational Linguistics, USA

Conference Webmasters:

Andrew Lampert, CSIRO ICT Centre, Australia
Brett Powley, Macquarie University, Australia

Conference Secretariat Coordinator:

Judy Potter, Well Done Events, Australia

Graphic Design:

Kathie Mason, Macquarie University, Australia

Advisory Committee:

Sandra Carberry, University of Delaware, USA
Eva Hajicova, Charles University, Czech Republic
Aravind K. Joshi, University of Pennsylvania, USA
Martin Kay, Stanford University, USA
Kathleen McCoy, University of Delaware, USA
Martha Palmer, University of Colorado, USA
Priscilla Rasmussen, Association for Computational Linguistics, USA
Benjamin Tsou, City University of Hong Kong, Hong Kong
Jun'ichi Tsujii, University of Tokyo, Japan

International Committee on Computational Linguistics (ICCL):

Igor Boguslavsky, Russian Academy of Sciences, Russia
Christian Boitet, GETA, CLIPS, IMAG, France
Nicoletta Calzolari, Istituto di Linguistica Computazionale – CNR, Italy
Eva Hajicova, Charles University, Czech Republic
Kolbjorn Heggstad
Chu-Ren Huang, Academia Sinica, Taiwan
Pierre Isabelle, National Research Council of Canada, Canada

Aravind K. Joshi, University of Pennsylvania, USA
Martin Kay, Stanford University, USA
Winfried Lenders, IKS-Universitaet Bonn, Germany
Makoto Nagao, Kyoto University, Japan
Sergei Nirenburg, University of Maryland Baltimore County, USA
Helmut Schnelle
Donia Scott, The Open University, UK
Petr Sgall, Charles University, Czech Republic
Hozumi Tanaka, Tokyo Institute of Technology, Japan
Jun'ichi Tsujii, University of Tokyo, Japan
Hans Uszkoreit, DFKI Saarbruecken, Germany
Hiroshi Wada
Yorick Wilks, University of Sheffield, UK

ACL Executive Committee:

Jun'ichi Tsujii, University of Tokyo, Japan
Mark Steedman, University of Edinburgh, UK
Bonnie Dorr, University of Maryland, USA
Kathleen McCoy, University of Delaware, USA
Dragomir Radev, University of Michigan, USA
Martha Palmer, University of Colorado, USA
Sandra Carberry, University of Delaware, USA
Walter Daelemans, University of Antwerp, Belgium
Keh-Yih Su, Behavior Design Corporation, Taiwan
Claire Cardie, Cornell University, USA

Program Committee

Chairs:

Claire Cardie, Cornell University, USA
Pierre Isabelle, National Research Council of Canada, Canada

Area Chairs:

Johan Bos, Università di Roma “La Sapienza”, Italy
Jason Chang, National Tsing Hua University, Taiwan
David Chiang, USC Information Sciences Institute, USA
Eva Hajicova, Charles University, Czech Republic
Chu-Ren Huang, Academia Sinica, Taiwan
Martin Kay, Stanford University, USA
Emiel Krahmer, Tilburg University, The Netherlands
Roland Kuhn, National Research Council of Canada, Canada
Lillian Lee, Cornell University, USA
Yuji Matsumoto, Nara Institute of Technology, Japan
Dan Moldovan, University of Texas, USA
Mark-Jan Nederhof, University of Groningen, The Netherlands
Hwee Tou Ng, National University of Singapore, Singapore
John Prager, IBM Watson Research Center, USA
Anoop Sarkar, Simon Fraser University, Canada
Donia Scott, The Open University, UK
Simone Teufel, University of Cambridge, UK
Benjamin Tsou, City University of Hong Kong, Hong Kong
ChengXiang Zhai, University of Illinois, USA
Ming Zhou, Microsoft Beijing, China

Program Committee Members:

Anne Abeille, Eugene Agichtein, Eneko Agirre, David Ahn, Lars Ahrenberg, Miguel Alonso Pardo, Rie Ando, Elisabeth Andre, Galen Andrew, Shlomo Argamon, Masayuki Asahara, Tania Avgustinova, Necip Fazil Ayan

Srinivas Bangalore, Regina Barzilay, Roberto Basili, Tilman Becker, S M Bendre, Stefano Bertolo, Pushpak Bhattacharya, Steffen Bickel, Daniel Bikel, Mikhail Bilenko, Philippe Blache, William Black, Constantinos Boulis, Thorsten Brants, Eric Breck, Sabine Buchholz, Razvan Bunescu, John Burger, Donna Byron

Chris Callison-Burch, Jaime Carbonell, Xavier Carreras, Vitor Carvalho, Nuria Castell, Frantisek Cermak, Joyce Chai, Soumen Chakrabarti, Ciprian Chelba, Hsin-Hsi Chen, John Chen, Keh-Jiann Chen, Kuang-hua Chen, Lee-Feng Chien, Yejin Choi, Tat-Seng Chua, Ken Church, Stephen Clark, James Clarke, Michael Collins, Matteo Contolini, Koby Crammer, Mathias Creutz, Silviu Cucerzan, James Curran, Krzysztof Czuba

Walter Daelemans, Ido Dagan, Hal Daume III, Renato DeMori, Barbara Di Eugenio, Mona Diab, Christy Doran, Mark Dras, Amit Dubey, Kevin Duh

Noemie Elhadad, Katrin Erk, Andrea Esuli, Yair Even-Zohar

Marcello Federico, Christiane Fellbaum, Radu Florian, George Forman, George Foster, Anette Frank, Robert Frank, Alex Fraser, Dayne Freitag, Sadaoki Furui

Evgeniy Gabrilovich, Jianfeng Gao, Eric Gaussier, Mark Gawron, Ruifang Ge, Effi Georgala, Mazin Gilbert, Roxana Girju, Oren Glickman, Jade Goldstein, Yifan Gong, Silke Goronzy, Cyril Goutte, Gregory Grefenstette, Ralph Grishman

Nizar Habash, Udo Hahn, Thomas Hain, Jan Hajic, Keith Hall, Sanda Harabagiu, Mary Harper, James Henderson, John Henderson, Erhard Hinrichs, Graeme Hirst, Barbora Hladka, Julia Hockenmaier, Veronique Hoste, Shu-Kai Hsieh, Fei Huang, Liang Huang

Nancy Ide, Kentaro Inui, Hitoshi Isahara, Hideki Isozaki, Abe Ittycheriah

Paul Jacobs, Nathalie Japkowicz, Donghong Ji, Rong Jin, Hongyan Jing, Howard Johnson, Michael Johnston, Rosie Jones, Jean-Claude Junqua

Kyo Kageura, Laura Kallmeyer, Nanda Kambhatla, Noriko Kando, Boris Katz, Asanee Kawtrakul, Martin Kay, Frank Keller, Rodger Kibble, Adam Kilgarriff, Tracy Hooloway King, Alexandra Kinyon, Katrin Kirchhoff, Dan Klein, Kevin Knight, Alistair Knott, Philipp Koehn, Moshe Koppel, Kimmo Koskenniemi, Taku Kudo, Peter Kuehnlein, Jonas Kuhn, Roland Kuhn, Shankar Kumar, Oren Kurland, Sadao Kurohashi, K L Kwok, Olivia Kwong

Tom Lai, Irene Langkilde-Geary, Philippe Langlais, Mirella Lapata, Alex Lascarides, Alberto Lavelli, Alon Lavie, Victor Lavrenko, Guy Lebanon, Gary (Geunbae) Lee, Jong-Hyeok Lee, Young-Suk Lee, Oliver Lemon, Alessandro Lenci, Piroska Lendvai, David Lewis, Hang Li, Mu Li, Xin Li, Liz Liddy, Chin-Yew Lin, Jimmy Lin, Ken Litkowski, Bing Liu, Beth Logan, Marketa Lopatkova, Adam Lopez, Yajuan Lu, Xiaoqiang Luo

Qing Ma, Bente Maegaard, Milind Mahajan, Steve Maiorano, Rob Malouf, Daniel Marcu, Mitch Marcus, Katja Markert, Lluis Marquez, Erwin Marsi, Yuji Matsumoto, John Maxwell, Andrew McCallum, Diana McCarthy, Kathleen McCoy, Ryan McDonald, Dan Melamed, Helen Meng, Wolfgang Menzel, Paola Merlo, Detmar Meurers, Adam Meyers, Rada Mihalcea, Natasa Milic-Frayling, Eleni Miltsakaki, Ruslan Mitkov, Mandar Mitra, Vibhu Mittal, Yusuke Miyao, Dunja Mladenic, Dan Moldovan, Monica Monachini, Christof Monz, Johanna Moore, Alessandro Moschitti, Isabelle Moulinier, Dragos Munteanu, Masaki Murata

Masaaki Nagata, Sobha L Nair, Hiroshi Nakagawa, Satoshi Nakamura, Srini Narayanan, Wuritu Nashun, Roberto Navigli, Hermann Ney, Vincent Ng, Patrick Nguyen, Nicolas Nicolov, Jian-Yun Nie, Kamal Nigam, Takashi Ninomiya, Malvina Nissim, Cheng Niu, Zheng-Yu Niu, Joakim Nivre, Eric Nyberg

Jon Oberlander, Franz Och, Stephan Oepen, Kemal Oflazer, Miles Osborne

Petr Pajas, Karel Pala, Martha Palmer, Jarmila Panevova, Bo Pang, Patrick Pantel, Fuchun Peng, Gerald Penn, Vladimqr Petkevic, Fabio Pianesi, Paul Piwek, Massimo Poesio, Patrice Pognan, Ana-Maria Popescu, Andrei Popescu-Belis, Richard Power, Sameer Pradhan, John Prange, Rashmi Prasad, Detlef Prescher, Laurent Prevot, Katharina Probst, Gabor Proszeky, Josef Psutka

Yan Qu

Owen Rambow, Alan Ramsay, Philip Resnik, Stefan Riezler, German Rigau, Hae-Chang Rim, Brian Roark, Rick Rose, Alex Rudnicky, Pavel Rychly

Carl Sable, Louisa Sadler, Kenji Sagae, Antonio Sanfilippo, Rajeev Sangal, Sunita Sarawagi, Giorgio Satta, Michael Schiehlen, Frank Schilder, Helmut Schmid, Hinrich Schuetze, Tanja Schultz, Holger Schwenk, Donia Scott, Jiri Semecky, Petr Sgall, Fei Sha, Vijay Shanker, Dipti M Sharma,

Libin Shen, Xiaodong Shi, Atsushi Shimojima, Luo Si, Advaith Siddharthan, Khalil Simaan, Man-Hung Siu, David Smith, Noah Smith, Harold Somers, Virach Sornlertlamvanich, Karen Sparck-Jones, Rohini Srihari, Manfred Stede, Mark Steedman, Amanda Stent, Suzanne Stevenson, Matthew Stone, Veselin Stoyanov, Carlo Strapparava, Tomek Strzalkowski, Jian Su, Keh-Yih Su, Maosong Sun, Mihai Surdeanu, Jun Suzuki, Marc Swerts

Hiroya Takamura, Marta Tatu, Thanaruk Theeramunkong, Mariet Theune, Christoph Tillmann, Takenobu Tokunaga, Andrew Tomkins, Kristina Toutanova, David Traum, Huihsin Tseng, Benjamin Tsou, Dan Tufis, Peter Turney

Kiyotaka Uchimoto, Nicola Ueffing, Takehito Utsuro

Antal van den Bosch, Josef van Genabith, Gertjan van Noord, Lucy Vanderwende, Hans vanHalteren, Ashish Venugopal, Peter Veprek, Stephan Vogel, Piek Vossen

Marilyn Walker, Shaojun Wang, Wei Wang, Xiaolong Wang, Ye-Yi Wang, Andy Way, Bonnie Webber, David Weir, Michael White, Ed Whittaker, Richard Wicentowski, Jan Wiebe, Yorick Wilks, Theresa Wilson, Shuly Wintner, Dekai Wu, Xiaoyuan Wu, Yunfang Wu

Fei Xia, Peng Xu

Z Zabokrtsky, Fabio Massimo Zanzotto, Daniel Zeman, Richard Zens, Hao Zhang, Tong Zhang, Yi Zhang, Ying Zhang, GuoDong Zhou, Jingbo Zhu, Imed Zitouni

Conference Program

Monday, 17 July 2006

09:00–09:30 Opening Session

Session 1A: Machine Translation I

09:30–10:00 *Combination of Arabic Preprocessing Schemes for Statistical Machine Translation*
Fatiha Sadat and Nizar Habash

10:00–10:30 *Going Beyond AER: An Extensive Analysis of Word Alignments and Their Impact on MT*
Necip Fazil Ayan and Bonnie J. Dorr

Session 1B: Topic Segmentation

09:30–10:00 *Unsupervised Topic Modelling for Multi-Party Spoken Discourse*
Matthew Purver, Konrad P. Körding, Thomas L. Griffiths and Joshua B. Tenenbaum

10:00–10:30 *Minimum Cut Model for Spoken Lecture Segmentation*
Igor Malioutov and Regina Barzilay

Session 1C: Coreference

09:30–10:00 *Bootstrapping Path-Based Pronoun Resolution*
Shane Bergsma and Dekang Lin

10:00–10:30 *Kernel-Based Pronoun Resolution with Structured Syntactic Knowledge*
Xiaofeng Yang, Jian Su and Chew Lim Tan

Session 1D: Grammars I

09:30–10:00 *A Finite-State Model of Human Sentence Processing*
Jihyun Park and Chris Brew

10:00–10:30 *Acceptability Prediction by Means of Grammaticality Quantification*
Philippe Blache, Barbara Hemforth and Stéphane Rauzy

10:30–11:00 Break

Monday, 17 July 2006 (continued)

Session 2A: Machine Translation II

11:00–11:30 *Discriminative Word Alignment with Conditional Random Fields*
Phil Blunsom and Trevor Cohn

11:30–12:00 *Named Entity Transliteration with Comparable Corpora*
Richard Sproat, Tao Tao and ChengXiang Zhai

12:00–12:30 *Extracting Parallel Sub-Sentential Fragments from Non-Parallel Corpora*
Dragos Stefan Munteanu and Daniel Marcu

Session 2B: Word Sense Disambiguation I

11:00–11:30 *Estimating Class Priors in Domain Adaptation for Word Sense Disambiguation*
Yee Seng Chan and Hwee Tou Ng

11:30–12:00 *Ensemble Methods for Unsupervised WSD*
Samuel Brody, Roberto Navigli and Mirella Lapata

12:00–12:30 *Meaningful Clustering of Senses Helps Boost Word Sense Disambiguation Performance*
Roberto Navigli

Session 2C: Information Extraction I

11:00–11:30 *Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations*
Patrick Pantel and Marco Pennacchiotti

11:30–12:00 *Modeling Commonality among Related Classes in Relation Extraction*
GuoDong Zhou, Jian Su and Min Zhang

12:00–12:30 *Relation Extraction Using Label Propagation Based Semi-Supervised Learning*
Jinxiu Chen, Donghong Ji, Chew Lim Tan and Zhengyu Niu

Session 2D: Grammars II

11:00–11:30 *Polarized Unification Grammars*
Sylvain Kahane

11:30–12:00 *Partially Specified Signatures: A Vehicle for Grammar Modularity*
Yael Cohen-Sygal and Shuly Wintner

12:00–12:30 *Morphology-Syntax Interface for Turkish LFG*
Özlem Çetinoğlu and Kemal Oflazer

12:30–14:00 Lunch

Monday, 17 July 2006 (continued)

Session 3A: Parsing I

14:00–14:30 *PCFGs with Syntactic and Prosodic Indicators of Speech Repairs*
John Hale, Izhak Shafran, Lisa Yung, Bonnie Dorr, Mary Harper, Anna Krasnyanskaya, Matthew Lease, Yang Liu, Brian Roark, Matthew Snover and Robin Stewart

14:30–15:00 *Dependency Parsing of Japanese Spoken Monologue Based on Clause Boundaries*
Tomohiro Ohno, Shigeki Matsubara, Hideki Kashioka, Takehiko Maruyama and Yasuyoshi Inagaki

15:00–15:30 *Trace Prediction and Recovery with Unlexicalized PCFGs and Slash Features*
Helmut Schmid

Session 3B: Dialogue I

14:00–14:30 *Learning More Effective Dialogue Strategies Using Limited Dialogue Move Features*
Matthew Frampton and Oliver Lemon

14:30–15:00 *Dependencies between Student State and Speech Recognition Problems in Spoken Tutoring Dialogues*
Mihai Rotaru and Diane J. Litman

15:00–15:30 *Learning the Structure of Task-Driven Human-Human Dialogs*
Srinivas Bangalore, Giuseppe Di Fabbrizio and Amanda Stent

Session 3C: Machine Learning Methods I

14:00–14:30 *Semi-Supervised Conditional Random Fields for Improved Sequence Segmentation and Labeling*
Feng Jiao, Shaojun Wang, Chi-Hoon Lee, Russell Greiner and Dale Schuurmans

14:30–15:00 *Training Conditional Random Fields with Multivariate Evaluation Measures*
Jun Suzuki, Erik McDermott and Hideki Isozaki

15:00–15:30 *Approximation Lasso Methods for Language Modeling*
Jianfeng Gao, Hisami Suzuki and Bin Yu

Session 3D: Applications I

14:00–14:30 *Automated Japanese Essay Scoring System based on Articles Written by Experts*
Tsunenori Ishioka and Masayuki Kameda

14:30–15:00 *A Feedback-Augmented Method for Detecting Errors in the Writing of Learners of English*
Ryo Nagata, Atsuo Kawai, Koichiro Morihiro and Naoki Isu

15:00–15:30 *Correcting ESL Errors Using Phrasal SMT Techniques*
Chris Brockett, William B. Dolan and Michael Gamon

15:30–16:00 Break

Monday, 17 July 2006 (continued)

Session 4A: Parsing II

16:00–16:30 *Graph Transformations in Data-Driven Dependency Parsing*
Jens Nilsson, Joakim Nivre and Johan Hall

Session 4B: Dialogue II

16:00–16:30 *Learning to Generate Naturalistic Utterances Using Reviews in Spoken Dialogue Systems*
Ryuichiro Higashinaka, Rashmi Prasad and Marilyn A. Walker

Session 4C: Linguistic Kinships

16:00–16:30 *Measuring Language Divergence by Intra-Lexical Comparison*
T. Mark Ellison and Simon Kirby

Session 4D: Applications II

16:00–16:30 *Enhancing Electronic Dictionaries with an Index Based on Associations*
Olivier Ferret and Michael Zock

16:30–17:30 ACL Lifetime Achievement Award

17:30–19:30 Poster Sessions

Tuesday, 18 July 2006

09:00–10:00 Invited Talk by Daniel Marcu: $Argmax$ *Search in Natural Language Processing*

Session 5A: Parsing III

10:00–10:30 *Guiding a Constraint Dependency Parser with Supertags*
Kilian Foth, Tomas By and Wolfgang Menzel

Session 5B: Lexical Issues I

10:00–10:30 *Efficient Unsupervised Discovery of Word Categories Using Symmetric Patterns and High Frequency Words*
Dmitry Davidov and Ari Rappoport

Session 5C: Summarization I

10:00–10:30 *Bayesian Query-Focused Summarization*
Hal Daumé III and Daniel Marcu

Session 5D: Semantics I

10:00–10:30 *Expressing Implicit Semantic Relations without Supervision*
Peter D. Turney

10:30–11:00 Break

Tuesday, 18 July 2006 (continued)

Session 6A: Parsing IV

11:00–11:30 *Hybrid Parsing: Using Probabilistic Models as Predictors for a Symbolic Parser*
Kilian A. Foth and Wolfgang Menzel

11:30–12:00 *Error Mining in Parsing Results*
Benoît Sagot and Éric de La Clergerie

12:00–12:30 *Reranking and Self-Training for Parser Adaptation*
David McClosky, Eugene Charniak and Mark Johnson

Session 6B: Lexical Issues II

11:00–11:30 *Automatic Classification of Verbs in Biomedical Texts*
Anna Korhonen, Yuval Krymolowski and Nigel Collier

11:30–12:00 *Selection of Effective Contextual Information for Automatic Synonym Acquisition*
Masato Hagiwara, Yasuhiro Ogawa and Katsuhiko Toyama

12:00–12:30 *Scaling Distributional Similarity to Large Corpora*
James Gorman and James R. Curran

Session 6C: Summarization II

11:00–11:30 *Extractive Summarization using Inter- and Intra- Event Relevance*
Wenjie Li, Mingli Wu, Qin Lu, Wei Xu and Chunfa Yuan

11:30–12:00 *Models for Sentence Compression: A Comparison across Domains, Training Requirements and Evaluation Measures*
James Clarke and Mirella Lapata

12:00–12:30 *A Bottom-Up Approach to Sentence Ordering for Multi-Document Summarization*
Danushka Bollegala, Naoaki Okazaki and Mitsuru Ishizuka

Session 6D: Semantics II

11:00–11:30 *Learning Event Durations from Event Descriptions*
Feng Pan, Rutu Mulkar and Jerry R. Hobbs

11:30–12:00 *Automatic Learning of Textual Entailments with Cross-Pair Similarities*
Fabio Massimo Zanzotto and Alessandro Moschitti

12:00–12:30 *An Improved Redundancy Elimination Algorithm for Underspecified Representations*
Alexander Koller and Stefan Thater

12:30–14:00 Lunch

Tuesday, 18 July 2006 (continued)

Session 7A: Parsing V

14:00–14:30 *Integrating Syntactic Priming into an Incremental Probabilistic Parser, with an Application to Psycholinguistic Modeling*
Amit Dubey, Frank Keller and Patrick Sturt

14:30–15:00 *A Fast, Accurate Deterministic Parser for Chinese*
Mengqiu Wang, Kenji Sagae and Teruko Mitamura

15:00–15:30 *Learning Accurate, Compact, and Interpretable Tree Annotation*
Slav Petrov, Leon Barrett, Romain Thibaux and Dan Klein

Session 7B: Word Sense Disambiguation II

14:00–14:30 *Semi-Supervised Learning of Partial Cognates Using Bilingual Bootstrapping*
Oana Frunza and Diana Inkpen

14:30–15:00 *Direct Word Sense Matching for Lexical Substitution*
Ido Dagan, Oren Glickman, Alfio Gliozzo, Efrat Marmorshtein and Carlo Strapparava

15:00–15:30 *An Equivalent Pseudoword Solution to Chinese Word Sense Disambiguation*
Zhimao Lu, Haifeng Wang, Jianmin Yao, Ting Liu and Sheng Li

Session 7C: Information Extraction II

14:00–14:30 *Improving the Scalability of Semi-Markov Conditional Random Fields for Named Entity Recognition*
Daisuke Okanohara, Yusuke Miyao, Yoshimasa Tsuruoka and Jun'ichi Tsujii

14:30–15:00 *Factorizing Complex Models: A Case Study in Mention Detection*
Radu Florian, Hongyan Jing, Nanda Kambhatla and Imed Zitouni

15:00–15:30 *Segment-Based Hidden Markov Models for Information Extraction*
Zhenmei Gu and Nick Cercone

Session 7D: Resources I

14:00–14:30 *A DOM Tree Alignment Model for Mining Parallel Data from the Web*
Lei Shi, Cheng Niu, Ming Zhou and Jianfeng Gao

14:30–15:00 *QuestionBank: Creating a Corpus of Parse-Annotated Questions*
John Judge, Aoife Cahill and Josef van Genabith

15:00–15:30 *Creating a CCGbank and a Wide-Coverage CCG Lexicon for German*
Julia Hockenmaier

15:30–16:00 Break

Tuesday, 18 July 2006 (continued)

Session 8A: Machine Translation III

16:00–16:30 *Improved Discriminative Bilingual Word Alignment*
Robert C. Moore, Wen-tau Yih and Andreas Bode

16:30–17:00 *Maximum Entropy Based Phrase Reordering Model for Statistical Machine Translation*
Deyi Xiong, Qun Liu and Shouxun Lin

17:00–17:30 *Distortion Models for Statistical Machine Translation*
Yaser Al-Onaizan and Kishore Papineni

Session 8B: Text Classification I

16:00–16:30 *A Study on Automatically Extracted Keywords in Text Categorization*
Anette Hulth and Beáta B. Megyesi

16:30–17:00 *A Comparison and Semi-Quantitative Analysis of Words and Character-Bigrams as Features in Chinese Text Categorization*
Jingyang Li, Maosong Sun and Xian Zhang

17:00–17:30 *Exploiting Comparable Corpora and Bilingual Dictionaries for Cross-Language Text Categorization*
Alfio Gliozzo and Carlo Strapparava

Session 8C: Machine Learning Methods II

16:00–16:30 *A Progressive Feature Selection Algorithm for Ultra Large Feature Spaces*
Qi Zhang, Fuliang Weng and Zhe Feng

16:30–17:00 *Annealing Structural Bias in Multilingual Weighted Grammar Induction*
Noah A. Smith and Jason Eisner

17:00–17:30 *Maximum Entropy Based Restoration of Arabic Diacritics*
Imed Zitouni, Jeffrey S. Sorensen and Ruhi Sarikaya

Session 8D: Information Retrieval I

16:00–16:30 *An Iterative Implicit Feedback Approach to Personalized Search*
Yuanhua Lv, Le Sun, Junlin Zhang, Jian-Yun Nie, Wan Chen and Wei Zhang

16:30–17:00 *The Effect of Translation Quality in MT-Based Cross-Language Information Retrieval*
Jiang Zhu and Haifeng Wang

17:00–17:30 *A Comparison of Document, Sentence, and Term Event Spaces*
Catherine Blake

17:30–19:30 Poster Sessions

Thursday, 20 July 2006

Session 9A: Best Asian Language Paper Nominee

09:00–09:30 *Tree-to-String Alignment Template for Statistical Machine Translation*
Yang Liu, Qun Liu and Shouxun Lin

Session 9B: Best Asian Language Paper Nominee

09:00–09:30 *Incorporating Speech Recognition Confidence into Discriminative Named Entity Recognition of Speech Data*
Katsuhito Sudoh, Hajime Tsukada and Hideki Isozaki

Session 9C: Best Asian Language Paper Nominee

09:00–09:30 *Exploiting Syntactic Patterns as Clues in Zero-Anaphora Resolution*
Ryu Iida, Kentaro Inui and Yuji Matsumoto

Session 9D: Best Asian Language Paper Nominee

09:00–09:30 *Self-Organizing n-gram Model for Automatic Word Spacing*
Seong-Bae Park, Yoon-Shik Tae and Se-Young Park

09:30–10:30 Asian Language Special Event: *Challenges in NLP: Some New Perspectives from the East*

10:30–11:00 Break

Thursday, 20 July 2006 (continued)

Session 10A: Asian Language Processing

11:00–11:30 *Concept Unification of Terms in Different Languages for IR*
Qing Li, Sung-Hyon Myaeng, Yun Jin and Bo-yeong Kang

11:30–12:00 *Word Alignment in English-Hindi Parallel Corpus Using Recency-Vector Approach: Some Studies*
Niladri Chatterjee and Saumya Agrawal

12:00–12:30 *Extracting Loanwords from Mongolian Corpora and Producing a Japanese-Mongolian Bilingual Dictionary*
Badam-Osor Khaltar, Atsushi Fujii and Tetsuya Ishikawa

Session 10B: Morphology and Word Segmentation

11:00–11:30 *An Unsupervised Morpheme-Based HMM for Hebrew Morphological Disambiguation*
Meni Adler and Michael Elhadad

11:30–12:00 *Contextual Dependencies in Unsupervised Word Segmentation*
Sharon Goldwater, Thomas L. Griffiths and Mark Johnson

12:00–12:30 *MAGEAD: A Morphological Analyzer and Generator for the Arabic Dialects*
Nizar Habash and Owen Rambow

Session 10C: Tagging and Chunking

11:00–11:30 *Noun Phrase Chunking in Hebrew: Influence of Lexical and Morphological Features*
Yoav Goldberg, Meni Adler and Michael Elhadad

11:30–12:00 *Multi-Tagging for Lexicalized-Grammar Parsing*
James R. Curran, Stephen Clark and David Vadas

12:00–12:30 *Guessing Parts-of-Speech of Unknown Words Using Global Information*
Tetsuji Nakagawa and Yuji Matsumoto

12:30–13:30 Lunch

13:30–14:30 ACL Business Meeting

Thursday, 20 July 2006 (continued)

Session 11A: Machine Translation IV

14:30–15:00 *A Clustered Global Phrase Reordering Model for Statistical Machine Translation*
Masaaki Nagata, Kuniko Saito, Kazuhide Yamamoto and Kazuteru Ohashi

15:00–15:30 *A Discriminative Global Training Algorithm for Statistical MT*
Christoph Tillmann and Tong Zhang

Session 11B: Speech

14:30–15:00 *Phoneme-to-Text Transcription System with an Infinite Vocabulary*
Shinsuke Mori, Daisuke Takuma and Gakuto Kurata

15:00–15:30 *Automatic Generation of Domain Models for Call-Centers from Noisy Transcriptions*
Shourya Roy and L Venkata Subramaniam

Session 11C: Discourse

14:30–15:00 *Proximity in Context: An Empirically Grounded Computational Model of Proximity for Processing Topological Spatial Expressions*
John D. Kelleher, Geert-Jan M. Kruijff and Fintan J. Costello

15:00–15:30 *Machine Learning of Temporal Relations*
Inderjeet Mani, Marc Verhagen, Ben Wellner, Chong Min Lee and James Pustejovsky

15:30–16:00 Break

Thursday, 20 July 2006 (continued)

Session 12A: Machine Translation V

16:00–16:30 *An End-to-End Discriminative Approach to Machine Translation*
Percy Liang, Alexandre Bouchard-Côté, Dan Klein and Ben Taskar

16:30–17:00 *Semi-Supervised Training for Statistical Word Alignment*
Alexander Fraser and Daniel Marcu

17:00–17:30 *Left-to-Right Target Generation for Hierarchical Phrase-Based Translation*
Taro Watanabe, Hajime Tsukada and Hideki Isozaki

Session 12B: Lexical Issues III

16:00–16:30 *You Can't Beat Frequency (Unless You Use Linguistic Knowledge) – A Qualitative Evaluation of Association Measures for Collocation and Term Extraction*
Joachim Wermter and Udo Hahn

16:30–17:00 *Ontologizing Semantic Relations*
Marco Pennacchiotti and Patrick Pantel

17:00–17:30 *Semantic Taxonomy Induction from Heterogenous Evidence*
Rion Snow, Daniel Jurafsky and Andrew Y. Ng

Session 12C: Information Extraction III

16:00–16:30 *Names and Similarities on the Web: Fact Extraction in the Fast Lane*
Marius Paşca, Dekang Lin, Jeffrey Bigham, Andrei Lifchits and Alpa Jain

16:30–17:00 *Weakly Supervised Named Entity Transliteration and Discovery from Multilingual Comparable Corpora*
Alexandre Klementiev and Dan Roth

17:00–17:30 *A Composite Kernel to Extract Relations between Entities with Both Flat and Structured Features*
Min Zhang, Jie Zhang, Jian Su and GuoDong Zhou

Friday, 21 July 2006

9:00–10:00 Invited Talk by Sally McConnell-Ginet: *Language, Gender and Sexuality: Do Bodies Always Matter?*

Session 13A: Parsing VI

10:00–10:30 *Japanese Dependency Parsing Using Co-Occurrence Information and a Combination of Case Elements*
Takeshi Abekawa and Manabu Okumura

Session 13B: Question Answering I

10:00–10:30 *Answer Extraction, Semantic Clustering, and Extractive Summarization for Clinical Question Answering*
Dina Demner-Fushman and Jimmy Lin

Session 13C: Semantics III

10:00–10:30 *Discovering Asymmetric Entailment Relations between Verbs Using Selectional Preferences*
Fabio Massimo Zanzotto, Marco Pennacchiotti and Maria Teresa Pazienza

Session 13D: Applications III

10:00–10:30 *Event Extraction in a Plot Advice Agent*
Harry Halpin and Johanna D. Moore

10:30–11:00 Break

Friday, 21 July 2006 (continued)

Session 14A: Parsing VII

11:00–11:30 *An All-Subtrees Approach to Unsupervised Parsing*
Rens Bod

11:30–12:00 *Advances in Discriminative Parsing*
Joseph Turian and I. Dan Melamed

12:00–12:30 *Prototype-Driven Grammar Induction*
Aria Haghighi and Dan Klein

Session 14B: Question Answering II

11:00–11:30 *Exploring Correlation of Dependency Relation Paths for Answer Extraction*
Dan Shen and Dietrich Klakow

11:30–12:00 *Question Answering with Lexical Chains Propagating Verb Arguments*
Adrian Novischi and Dan Moldovan

12:00–12:30 *Methods for Using Textual Entailment in Open-Domain Question Answering*
Sanda Harabagiu and Andrew Hickl

Session 14C: Semantics IV

11:00–11:30 *Using String-Kernels for Learning Semantic Parsers*
Rohit J. Kate and Raymond J. Mooney

11:30–12:00 *A Bootstrapping Approach to Unsupervised Detection of Cue Phrase Variants*
Rashid M. Abdalla and Simone Teufel

12:00–12:30 *Semantic Role Labeling via FrameNet, VerbNet and PropBank*
Ana-Maria Giuglea and Alessandro Moschitti

Session 14D: Resources II

11:00–11:30 *Multilingual Legal Terminology on the Jibiki Platform: The LexALP Project*
Gilles Sérasset, Francis Brunet-Manquat and Elena Chiocchetti

11:30–12:00 *Leveraging Reusability: Cost-Effective Lexical Acquisition for Large-Scale Ontology Translation*
G. Craig Murray, Bonnie J. Dorr, Jimmy Lin, Jan Hajič and Pavel Pecina

12:00–12:30 *Accurate Collocation Extraction Using a Multilingual Parser*
Violeta Seretan and Eric Wehrli

12:30–14:00 Lunch

Friday, 21 July 2006 (continued)

Session 15A: Machine Translation VI

14:00–14:30 *Scalable Inference and Training of Context-Rich Syntactic Translation Models*
Michel Galley, Jonathan Graehl, Kevin Knight, Daniel Marcu, Steve DeNeefe, Wei Wang and Ignacio Thayer

14:30–15:00 *Modelling Lexical Redundancy for Machine Translation*
David Talbot and Miles Osborne

15:00–15:30 *Empirical Lower Bounds on the Complexity of Translational Equivalence*
Benjamin Wellington, Sonjia Waxmonsky and I. Dan Melamed

Session 15B: Language Modeling

14:00–14:30 *A Hierarchical Bayesian Language Model Based On Pitman-Yor Processes*
Yee Whye Teh

14:30–15:00 *A Phonetic-Based Approach to Chinese Chat Text Normalization*
Yunqing Xia, Kam-Fai Wong and Wenjie Li

15:00–15:30 *Discriminative Pruning of Language Models for Chinese Word Segmentation*
Jianfeng Li, Haifeng Wang, Dengjun Ren and Guohua Li

Session 15C: Information Retrieval II

14:00–14:30 *Novel Association Measures Using Web Search with Double Checking*
Hsin-Hsi Chen, Ming-Shun Lin and Yu-Chuan Wei

14:30–15:00 *Semantic Retrieval for the Accurate Identification of Relational Concepts in Massive Textbases*
Yusuke Miyao, Tomoko Ohta, Katsuya Masuda, Yoshimasa Tsuruoka, Kazuhiro Yoshida, Takashi Ninomiya and Jun'ichi Tsujii

15:00–15:30 *Exploring Distributional Similarity Based Models for Query Spelling Correction*
Mu Li, Muhua Zhu, Yang Zhang and Ming Zhou

Session 15D: Generation I

14:00–14:30 *Robust PCFG-Based Generation Using Automatically Acquired LFG Approximations*
Aoife Cahill and Josef van Genabith

14:30–15:00 *Incremental Generation of Spatial Referring Expressions in Situated Dialog*
John D. Kelleher and Geert-Jan M. Kruijff

15:00–15:30 *Learning to Predict Case Markers in Japanese*
Hisami Suzuki and Kristina Toutanova

15:30–16:00 Break

Friday, 21 July 2006 (continued)

Session 16A: Text Classification II

16:00–16:30 *Are These Documents Written from Different Perspectives? A Test of Different Perspectives Based on Statistical Distribution Divergence*
Wei-Hao Lin and Alexander Hauptmann

16:30–17:00 *Word Sense and Subjectivity*
Janyce Wiebe and Rada Mihalcea

Session 16B: Question Answering III

16:00–16:30 *Improving QA Accuracy by Question Inversion*
John Prager, Pablo Duboue and Jennifer Chu-Carroll

16:30–17:00 *Reranking Answers for Definitional QA Using Language Modeling*
Yi Chen, Ming Zhou and Shilong Wang

Session 16C: Grammars III

16:00–16:30 *Highly Constrained Unification Grammars*
Daniel Feinstein and Shuly Wintner

16:30–17:00 *A Polynomial Parsing Algorithm for the Topological Model: Synchronizing Constituent and Dependency Grammars, Illustrated by German Word Order Phenomena*
Kim Gerdes and Sylvain Kahane

Session 16D: Generation II

16:00–16:30 *Stochastic Language Generation Using WIDL-Expressions and its Application in Machine Translation and Summarization*
Radu Soricut and Daniel Marcu

16:30–17:00 *Learning to Say It Well: Reranking Realizations by Predicted Synthesis Quality*
Crystal Nakatsu and Michael White

17:00–17:30 Closing Session