ACL-05

43rd Annual Meeting of the Association for Computational Linguistics

Proceedings of the Conference

25-30 June 2005
University of Michigan

Production and Manufacturing by
Omnipress Inc.
Post Office Box 7214
Madison, WI 53707-7214

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
75 Paterson Street, Suite 9
New Brunswick, NJ 08901
USA
Tel: +1-732-342-9100
Fax: +1-732-342-9339
acl@aclweb.org

Preface: General Chair

When I asked what General Chairs are supposed to worry about, the main advice I got was this: Make sure the Program Co-Chairs are in the same time zone. Well, what's fun about that? Between Istanbul, Singapore, and Los Angeles, we could easily solve problems in real time; by the time Kemal arrived at the office after breakfast, Hwee Tou was just back from lunch, and I was done carrying out the late-night raid on my own refrigerator. No problem.

I'd like to start by thanking everyone who submitted research work to ACL-05. I'd especially like to thank researchers new to the field – this is a great time to be in computational linguistics. Excellent research is one of the Two Critical Ingredients of a successful ACL conference.

Program Co-Chairs **Hwee Tou Ng** and **Kemal Oflazer** deserve our gratitude for putting an immense amount of work into the main session program. They and the Area Chairs got a large number of submissions this year, and the program is diverse and exciting. Thanks also to **Erika Barragan-Nunez** for arranging the program committee meeting in California.

Stefan Riezler assembled a program of five excellent tutorials to begin the meeting, and **Mirella Lapata** organized the workshop program, assisted by **Mark Dras**, **Mary Harper**, **Dan Klein**, and **Shuly Winter**. **Masaaki Nagata** and **Ted Pedersen** put together a high-quality demo session, including software systems from all over the world.

Jason Eisner and **Philipp Koehn** put in a tremendous amount of thought, effort, and persistence into publications. Each time ACL doubles the number of papers, the work way more than doubles. **Mark Johnson**, as sponsorship chair, requested that the money be shown to him (and it was!), so thanks very much to the sponsors, and to Mark. **Richard Wicentowski** took on two chair roles – exhibits and publicity – the latter of which included writing the useful ACL-05 newsletters forwarded by ubiquifamous **Priscilla Rasmussen**.

Regina Barzilay, **Chris Callison-Burch** and **Stephen Wan** organized the Student Research Workshop (and thanks again to all the students who submitted their research). **Richard Power** graciously agreed to do pre-submission mentoring for authors. The ACL Executive Committee provided help on a number of issues and responded quickly to questions – thank you, **Martha Palmer**, **Jun'ichi Tsujii**, **Mark Steedman**, **Kathy McCoy**, **Sandee Carberry**, **Johanna Moore**, **Priscilla Rasmussen**, **Annie Zaenen**, **Walter Daelemans**, and **Keh-Yih Su**.

Dragomir Radev went far beyond the call of duty as Local Arrangements Chair. He raised and solved lots of strategic issues, followed up on every wire and cable, and cajoled other ACL chairs into solving important problems fast. I believe he may even be responsible for the weather and for making sure your luggage arrived on the same day you did. Thanks to the whole local team: **Rich Thomason**, **Steve Abney**, **Joyce Chai**, **San Duanmu**, **Kurt Godden**, **Acrisio Pires**, **Martha Pollack**, **Keith van der Linden**, **Rick Lewis**, **Sara Schwartz**, and **Bill Vlisides**, and to **James Sweeney**, who served as the conference webmaster. On behalf of Dragomir, please let me thank the University of Michigan's School of Information, Department of Electrical Engineering and Computer Science, and Department of Linguistics for their support. Dragomir also arranged the banquet at the Henry Ford Museum, where ACL President **Martha Palmer** will no doubt make an excellent speech – that's of course the Other Critical Ingredient of a successful ACL.

Finally, I'd also like to thank all the other folks who helped create ACL-05, including student volunteers, exhibitors, tutorialists, and everyone else not listed here.

To ACL attendees: thanks for coming, and please have a good conference!

Kevin Knight
ACL-05 General Chair
May 9, 2005

Preface: Program Co-Chairs

Exciting research in computational linguistics is being pursued vigorously all over the world. This year, we received a record number of 423 submissions. The program committee accepted 77 papers, for an acceptance rate of 18%, continuing the tradition of the annual ACL conference as being one of the most competitive and selective conferences. Of the accepted papers, 42 are from North America, 18 from Europe and the Middle East, and 17 from Asia and Australia.

We would like to express our heartfelt gratitude to all the authors who submitted their papers, to the 231 program committee members who worked tirelessly to review all submissions, and to the ten Area Chairs who oversaw the review process, collated the reviews, led discussions on papers with conflicting reviews, and solicited additional reviews for controversial papers. The Program Committee Co-Chairs and the area chairs then met for two days at the program committee meeting held at USC/ISI to select the final set of accepted papers. We would like to thank **Kevin Knight**, the General Conference Chair, who made available USC/ISI as the meeting venue, and his assistant **Erika Barragan-Nunez** who took care of the meeting arrangements and logistics.

The ACL-05 main program lasts three days, and includes plenary sessions, three parallel paper sessions, demo and poster sessions, and the student research workshop. We are grateful to Professor **Justine Cassell** (Northwestern University) and Professor **Michael Jordan** (University of California, Berkeley) who have kindly accepted our invitation to present invited talks at the conference.

The ACL-05 conference will also feature the ACL Lifetime Achievement Award. This prestigious award is presented to a most distinguished researcher for his or her pioneering work in computational linguistics. Past distinguished recipients of this award are **Aravind Joshi**, **Makoto Nagao**, and **Karen Spärck-Jones**. The recipient of this award in 2005 will be announced at a special plenary session at ACL-05, followed by a special lecture by the award recipient. ACL-05 will also continue the tradition of presenting the Best Paper Award to an outstanding paper. This award will be announced in the plenary session at the end of the conference.

A conference like ACL would not succeed without the many volunteers who offer their generous help. We deeply appreciate the advice and support of **Kevin Knight**, General Conference Chair, **Dragomir Radev**, Local Arrangements Chair, and the Local Arrangements Committee. We are also grateful to the ACL Executive Committee for their guidance, and **Walter Daelemans** and **Marilyn Walker**, ACL-04 Program Co-Chairs, for sharing their experience. We would also like to thank **Jason Eisner** and **Philipp Koehn**, Publication Co-Chairs, for putting together the proceedings of this conference.

We wish you an enjoyable time at ACL-05!

Hwee Tou Ng and Kemal Oflazer
ACL-05 Program Co-Chairs
May 12, 2005

Organizers

General Conference Chair:
Kevin Knight, University of Southern California, USA

Program Chairs:
Hwee Tou Ng, National University of Singapore, Singapore
Kemal Oflazer, Sabancı University, Turkey

Tutorial Chair:
Stefan Riezler, PARC, USA

Workshop Chair:
Mirella Lapata, University of Edinburgh, UK

Workshop Committee:
Mark Dras, Macquarie University, Australia
Mary Harper, NSF and Purdue University, USA
Dan Klein, University of California, Berkeley, USA
Shuly Winter, University of Haifa, Israel

Demo Chairs:
Masaaki Nagata, ATR Labs, USA
Ted Pedersen, University of Minnesota, USA

Publication Chairs:
Jason Eisner, Johns Hopkins University, USA
Philipp Koehn, University of Edinburgh, UK

Sponsorshop Chair:
Mark Johnson, Brown University, USA

Exhibits and Publicity Chair:
Richard Wicentowski, Swarthmore College, USA

Student Research Workshop:
Regina Barzilay, Massachusetts Institute of Technology, USA
Chris Callison-Burch, University of Edinburgh, UK
Stephen Wan, Macquarie University, Australia

Local Organization Chair:
Dragomir Radev, University of Michigan, USA

Local Organization Committee:
Steve Abney, University of Michigan, USA
Joyce Chai, Michigan State University, USA

San Duanmu, University of Michigan, USA
Kurt Godden, General Motors Research, USA
Rick Lewis, University of Michigan, USA
Keith van der Linden, Calvin College, USA
Acrisio Pires, University of Michigan, USA
Martha Pollack, University of Michigan, USA
Sara Schwartz, University of Michigan, USA
Rich Thomason , University of Michigan, USA
Bill Vlisides, University of Michigan, USA

Conference Webmaster:

James Sweeney, University of Michigan, USA

ACL Executive Committee:

Martha Palmer, ACL president, University of Pennsylvania, USA
Jun'ichi Tsujii, University of Tokyo
Mark Steedman, University of Edinburgh, UK
Kathy McCoy, University of Delaware, USA
Sandee Carberry, University of Delaware, USA
Johanna Moore, University of Edinburgh, UK
Priscilla Rasmussen, ACL, USA
Annie Zaenen, PARC, USA
Walter Daelemans, University of Antwerp, Belgium
Keh-Yih Su, Behavior Design Corporation, Taiwan

Program Committee

Chairs

Hwee Tou Ng, National University of Singapore, Singapore
Kemal Oflazer, Sabancı University, Turkey

Area Chairs

Michael Collins, Massachusetts Institute of Technology, USA
Marti Hearst, University of California, Berkeley, USA
Hang Li, Microsoft Research, China
Chin-Yew Lin, University of Southern California, USA
Yuji Matsumoto, Nara Institute of Technology, Japan
Diana McCarthy, University of Sussex, UK
Hermann Ney, RWTH Aachen, Germany
Gerald Penn, University of Toronto, Canada
Brian Roark, Oregon Health and Science University, USA
Michael Strube, EML Research, Germany

Program Committee Members

Steve Abney (University of Michigan, USA), Eneko Agirre (University of the Basque Country, Spain), Yasemin Altun (Toyota Technological Institute at Chicago, USA), Rie Ando (IBM Research, USA), Chinatsu Aone (SRA International, USA), Masayuki Asahara (NAIST, Japan)

Tim Baldwin (University of Melbourne, Australia), Srinivas Bangalore (AT&T Research, USA), Regina Barzilay (Massachusetts Institute of Technology, USA), Dan Bikel (IBM Research, USA), Jeff Bilmes (University of Washington, USA), Steven Bird (University of Melbourne, Australia), Johan Bos (University of Edinburgh, UK), Antal van den Bosch (Tilburg University, The Netherlands), Thorsten Brants (Google, USA), Chris Brew (Ohio State University, USA), Ted Briscoe (University of Cambridge, UK), Bill Byrne (University of Cambridge, UK), Donna Byron (Ohio State University, USA)

Mary Elaine Califf (Illinois State University, USA), Jean Carletta (University of Edinburgh, UK), Bob Carpenter (Alias-i, USA), John Carroll (University of Sussex, UK), Francisco Casacuberta (Polytechnic University of Valencia, Spain), Justine Cassell (Northwestern University, USA), Nick Cercone (Dalhousie University, Canada), Mauro Cettolo (ITC-IRST, Italy), Eugene Charniak (Brown University, USA), Ciprian Chelba (Microsoft Research, USA), Aitao Chen (Yahoo, USA), Hsin-Hsi Chen (National Taiwan University, Taiwan), Stanley Chen (IBM Research, USA), David Chiang (University of Maryland, USA), Jennifer Chu-Carroll (IBM Research, USA), Tat-Seng Chua (National University of Singapore, Singapore), Grace Chung (CNRI, USA), Ken Church (Microsoft Research, USA), Fabio Ciravegna (University of Sheffield, UK), Alexander Clark (University of London, UK), Stephen Clark (University of Oxford, UK), William Cohen (Carnegie Mellon University, USA), Mark Craven (University of Wisconsin, USA), Richard Crouch (PARC, USA), James Curran (University of Sydney, Australia)

Jan Daciuk (Gdansk University of Technology, Poland), Walter Daelemans (University of Antwerp,

Belgium), Hal Daumé III (University of Southern California, USA), Barbara Di Eugenio (University of Illinois at Chicago, USA), Mona Diab (Stanford University, USA), Alexandre Dikovsky (University of Nantes, France), Bonnie Dorr (University of Maryland, USA)

Phil Edmonds (Sharp Laboratories of Europe Ltd, UK), Jason Eisner (Johns Hopkins University, USA), Stefan Evert (University of Osnabrück, Germany)

Marcello Federico (ITC-IRST, Italy), Christiane Fellbaum (Princeton University, USA), Radu Florian (IBM Research, USA), Eric Fosler-Lussier (Ohio State University, USA), George Foster (NRC Institute for Information Technology, Canada), Dayne Freitag (Carnegie Mellon University, USA), Jun'ichi Fukumoto (Ritsumeikan University, Japan), Pascale Fung (Hong Kong University of Science and Technology, Hong Kong)

Michel Galley (Columbia University, USA), Jianfeng Gao (Microsoft Research, China), Yuqing Gao (IBM Research, USA), Daniel Gildea (University of Rochester, USA), Jonathan Ginzburg (King's College London, UK), Jade Goldstein (Department of Defense, USA), Fernando Gomez (University of Central Florida, USA), Joshua Goodman (Microsoft Research, USA)

Kadri Hacioglu (University of Colorado, USA), Udo Hahn (University of Freiburg, Germany), Jan Hajic (Charles University, Czech Republic), Keith Hall (Johns Hopkins University, USA), Hans van Halteren (Radboud University Nijmegen, The Netherlands), Sanda Harabagiu (University of Texas at Dallas, USA), Marti Hearst (University of California, Berkeley, USA), James Henderson (University of Edinburgh, UK), Ulf Hermjakob (University of Southern California, USA), Don Hindle (Primus Knowledge Systems, USA), Graeme Hirst (University of Toronto, Canada), Julia Hockenmaier (University of Pennsylvania, USA), Eduard Hovy (University of Southern California, USA), Rebecca Hwa (University of Pittsburgh, USA)

Pierre Isabelle (NRC Institute for Information Technology, Canada), Shun Ishizaki (Keio University, Japan), Hideki Isozaki (NTT Communication Science Laboratories, Japan), Abraham Ittycheriah (IBM Research, USA)

Gerhard Jäger (University of Bielefeld, Germany), Nick Jakobi (Algorithmix, UK), Martin Jansche (Columbia University, USA), Mark Johnson (Brown University, USA), Dan Jurafsky (Stanford University, USA)

Laura Kallmeyer (University of Paris 7, France), Min-Yen Kan (National University of Singapore, Singapore), Ron Kaplan (PARC, USA), Lauri Karttunen (PARC, USA), Tsuneaki Kato (University of Tokyo, Japan), Andrew Kehler (University of California, San Diego, USA), Frank Keller (University of Edinburgh, UK), Adam Kilgarriff (Lexicography MasterClass Ltd, UK), Katrin Kirchhoff (University of Washington, USA), Esther Klabbers (Oregon Health and Science University, USA), Dietrich Klakow (Saarland University, Germany), Dan Klein (University of California, Berkeley, USA), Rob Koeling (University of Sussex, UK), Philipp Koehn (University of Edinburgh, UK), Alexander Koller (Saarland University, Germany), Anna Korhonen (University of Cambridge, UK), Kimmo Koskenniemi (University of Helsinki, Finland), Geert-Jan Kruijff (Saarland University, Germany), Taku Kudo (NTT, Japan), Jonas Kuhn (University

of Texas - Austin, USA), Shankar Kumar (Johns Hopkins University, USA), Sadao Kurohashi (University of Tokyo, Japan)

Geunbae Lee (POSTECH, South Korea), Jong-Hyeok Lee (POSTECH, South Korea), Gina Levow (University of Chicago, USA), Roger Levy (Stanford University, USA), Elizabeth Liddy (Syracuse University, USA), Marc Light (University of Iowa, USA), Dekang Lin (University of Alberta, Canada), Jimmy Lin (University of Maryland, USA), Diane Litman (University of Pittsburgh, USA), Xiaoquiang Luo (IBM Research, USA), Elliott Macklovitch (University of Montreal, Canada), Bernardo Magnini (ITC-IRST, Italy), Rob Malouf (San Diego State University, USA), Lidia Mangu (IBM Research, USA), Inderjeet Mani (Georgetown University, USA), Stanford University Manning Chris (USA), Daniel Marcu (University of Southern California, USA), Jose B. Marino (Technical University of Catalonia, Spain), Yuji Matsumoto (Nara Institute of Technology, Japan), Andrew McCallum (University of Massachusetts Amherst, USA), Diana McCarthy (University of Sussex, UK), Kathy McKeown (Columbia University, USA), Dan Melamed (New York University, USA), Wolfgang Menzel (University of Hamburg, Germany), Paola Merlo (University of Geneva, Switzerland), Detmars Meurers (Ohio State University, USA), Rada Mihalcea (University of North Texas, USA), Marie-France Moens (Catholic University of Leuven, Belgium), Bob Moore (Microsoft Research, Microsoft), Johanna Moore (University of Edinburgh, UK)

Masaaki Nagata (NTT, Japan), Preslav Nakov (University of California, Berkeley, USA), Mark-Jan Nederhof (University of Groningen, The Netherlands), Vincent Ng (University of Texas at Dallas, USA), Grace Ngai (Hong Kong Polytechnic University, Hong Kong), Tadashi Nomoto (National Institute of Japanese Literature, Japan), Gertjan van Noord (University of Groningen, The Netherlands)

Doug Oard (University of Maryland, USA), Franz Och (Google Research, USA), Miles Osborne (University of Edinburgh, UK), Mari Ostendorf (University of Washington, USA), Paul Over (NIST, USA)

Tim Paek (Microsoft Research, USA), Martha Palmer (University of Pennsylvania, USA), Patrick Pantel (University of Southern California, USA), Kishore Papineni (IBM Research, USA), Seong-Bae Park (Seoul National University, Korea), Catherine Pelachaud (University of Paris 8, France), Gerald Penn (University of Toronto, Canada), Fernando Pereira (University of Pennsylvania, USA), Massimo Poesio (University of Essex, UK), John Prager (IBM Research, USA)

Owen Rambow (Columbia University, USA), Deepak Ravichandran (University of Southern California, USA), Ehud Reiter (University of Aberdeen, UK), Norbert Reithinger (DFKI, Germany), Giuseppe Riccardi (AT&T Research, USA), Stefan Riezler (PARC, USA), German Rigau (University of the Basque Country, Spain), Ellen Riloff (University of Utah, USA), Hae-Chang Rim (Korea University, South Korea), Brian Roark (Oregon Health and Science University, USA), James Rogers (Earlham College, USA), Mats Rooth (Cornell University, USA), Barbara Rosario (University of California, Berkeley, USA), Dan Roth (University of Illinois at Urbana-

Champaign, USA), Salim Roukos (IBM Research, USA)

Horacio Saggion (University of Sheffield, UK), Mehran Sahami (Stanford University and Google, USA), Murat Saraclar (Boğaziçi University, Turkey), Anoop Sarkar (Simon Fraser University, Canada), Yutaka Sasaki (ATR, Japan), Hinrich Schütze (University of Stuttgart, Germany), Ralf Schlueter (RWTH Aachen, Germany), Sabine Schulte im Walde (Saarland University, Germany), Donia Scott (Open University, UK), Satoshi Sekine (New York University, USA), Zak Shafran (Johns Hopkins University, USA), Stuart Shieber (Harvard University, USA), Elizabeth Shriberg (SRI and ICSI, USA), Candy Sidner (MERL, USA), Khalil Sima'an (University of Amsterdam, The Netherlands), Karen Spärck Jones (Cambridge University, UK), Richard Sproat (University of Illinois at Urbana-Champaign, USA), Edward Stabler (UCLA, USA), Manfred Stede (University of Potsdam, Germany), Mark Steedman (University of Edinburgh, UK), Mark Stevenson (University of Sheffield, UK), Suzanne Stevenson (University of Toronto, Canada), Matthew Stone (Rutgers University, USA), Tomek Strzalkowski (SUNY, Albany, USA), Jian Su (Institute for Infocomm Research, Singapore), Keh-Yih Su (Behavior Design Corporation, Taiwan), Eiichiro Sumita (ATR, Japan)

Simone Teufel (Cambridge University, UK), Eric Tjong Kim Sang (University of Antwerp, Belgium), Kentaro Torisawa (Japan Advanced Institute of Science Technology, Japan), Evelyne Tzoukermann (StreamSage, USA)

Dimitra Vergyri (SRI International, USA), Enrique Vidal (Polytechnic University of Valencia, Spain), Aline Villavicencio (University of Essex, UK), Stephan Vogel (Carnegie Mellon University, USA), Ellen Voorhees (NIST, USA)

Wen Wang (SRI International, USA), XingLong Wang (University of Sussex, UK), Bonnie Webber (University of Edinburgh, UK), Julie Weeds (University of Sussex, UK), Janyce Wiebe (University of Pittsburgh, USA), Ross Wilkinson (CSIRO, Australia), Shuly Wintner (University of Haifa, Israel), Dekai Wu (Hong Kong University of Science and Technology, Hong Kong)

Peng Xu (Johns Hopkins University, USA)

Christopher Yang (The Chinese University of Hong Kong, Hong Kong)

Richard Zens (RWTH Aachen, Germany), ChengXiang Zhai (University of Illinois at Urbana-Champaign, USA), GuoDong Zhou (Institute for Infocomm Research, Singapore)

Table of Contents

Conference Program

Sunday, June 26, 2005

8:45–9:00 Opening

9:00–10:00 Invited Talk by Justine Cassell

10:00–10:30 Break

Session M1R: Machine Learning and Statistical Models

10:30–11:00 *A High-Performance Semi-Supervised Learning Method for Text Chunking*
Rie Ando and Tong Zhang

11:00–11:30 *Scaling Conditional Random Fields Using Error-Correcting Codes*
Trevor Cohn, Andrew Smith and Miles Osborne

11:30–12:00 *Logarithmic Opinion Pools for Conditional Random Fields*
Andrew Smith, Trevor Cohn and Miles Osborne

Session M1M: Word Sense Disambiguation

10:30–11:00 *Supersense Tagging of Unknown Nouns using Semantic Similarity*
James Curran

11:00–11:30 *Learning Semantic Classes for Word Sense Disambiguation*
Upali Sathyajith Kohomban and Wee Sun Lee

11:30–12:00 *The Role of Semantic Roles in Disambiguating Verb Senses*
Hoa Trang Dang and Martha Palmer

Sunday, June 26, 2005 (continued)

Session M1B: Generation

10:30–11:00 *Aggregation Improves Learning: Experiments in Natural Language Generation for Intelligent Tutoring Systems*
Barbara Di Eugenio, Davide Fossati, Dan Yu, Susan Haller and Michael Glass

11:00–11:30 *Empirically-based Control of Natural Language Generation*
Daniel S. Paiva and Roger Evans

11:30–12:00 *Towards Developing Generation Algorithms for Text-to-Text Applications*
Radu Soricut and Daniel Marcu

12:00–1:30 Lunch

Session M2R: Parsing

1:30–2:00 *Probabilistic CFG with Latent Annotations*
Takuya Matsuzaki, Yusuke Miyao and Jun'ichi Tsujii

2:00–2:30 *Probabilistic Disambiguation Models for Wide-Coverage HPSG Parsing*
Yusuke Miyao and Jun'ichi Tsujii

2:30–3:00 *Online Large-Margin Training of Dependency Parsers*
Ryan McDonald, Koby Crammer and Fernando Pereira

3:00–3:30 *Pseudo-Projective Dependency Parsing*
Joakim Nivre and Jens Nilsson

Sunday, June 26, 2005 (continued)

Session M2M: Semantics

1:30–2:00 *The Distributional Inclusion Hypotheses and Lexical Entailment*
Maayan Geffet and Ido Dagan

2:00–2:30 *Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales*
Bo Pang and Lillian Lee

2:30–3:00 *Inducing Ontological Co-occurrence Vectors*
Patrick Pantel

3:00–3:30 *Extracting Semantic Orientations of Words using Spin Model*
Hiroya Takamura, Takashi Inui and Manabu Okumura

Session M2B: Discourse

1:30–2:00 *Modeling Local Coherence: An Entity-Based Approach*
Regina Barzilay and Mirella Lapata

2:00–2:30 *Modelling the Substitutability of Discourse Connectives*
Ben Hutchinson

2:30–3:00 *Machine Learning for Coreference Resolution: From Local Classification to Global Ranking*
Vincent Ng

3:00–3:30 *Improving Pronoun Resolution Using Statistics-Based Semantic Compatibility Information*
Xiaofeng Yang, Jian Su and Chew Lim Tan

3:30–4:00 Break

Sunday, June 26, 2005 (continued)

Session M3R: Parsing

4:00–4:30 *Coarse-to-Fine n-Best Parsing and MaxEnt Discriminative Reranking*
Eugene Charniak and Mark Johnson

4:30–5:00 *Data-Defined Kernels for Parse Reranking Derived from Probabilistic Models*
James Henderson and Ivan Titov

5:00–5:30 *Boosting-based Parse Reranking with Subtree Features*
Taku Kudo, Jun Suzuki and Hideki Isozaki

5:30–6:00 *Automatic Measurement of Syntactic Development in Child Language*
Kenji Sagae, Alon Lavie and Brian MacWhinney

Session M3M: Question Answering

4:00–4:30 *Experiments with Interactive Question-Answering*
Sanda Harabagiu, Andrew Hickl, John Lehmann and Dan Moldovan

4:30–5:00 *Question Answering as Question-Biased Term Extraction: A New Approach toward Multilingual QA*
Yutaka Sasaki

Session M3B: Discourse and Dialogue

4:00–4:30 *Exploring and Exploiting the Limited Utility of Captions in Recognizing Intention in Information Graphics*
Stephanie Elzer, Sandra Carberry, Daniel Chester, Seniz Demir, Nancy Green, Ingrid Zukerman and Keith Trnka

4:30–5:00 *Scaling up from Dialogue to Multilogue: Some Principles and Benchmarks*
Jonathan Ginzburg and Raquel Fernández

5:00–5:30 *Implications for Generating Clarification Requests in Task-Oriented Dialogues*
Verena Rieser and Johanna Moore

5:30–6:00 *Towards Finding and Fixing Fragments - Using ML to Identify Non-Sentential Utterances and their Antecedents in Multi-Party Dialogue*
David Schlangen

Monday, June 27, 2005

Session M4R: Machine Translation

9:00–9:30 *Scaling Phrase-Based Statistical Machine Translation to Larger Corpora and Longer Phrases*
Chris Callison-Burch, Colin Bannard and Josh Schroeder

9:30–10:00 *A Hierarchical Phrase-Based Model for Statistical Machine Translation*
David Chiang

10:00–10:30 *Dependency Treelet Translation: Syntactically Informed Phrasal SMT*
Chris Quirk, Arul Menezes and Colin Cherry

Session M4M: Summarization

9:00–9:30 *QARLA: A Framework for the Evaluation of Text Summarization Systems*
Enrique Amigó, Julio Gonzalo, Anselmo Peñas and Felisa Verdejo

9:30–10:00 *Supervised and Unsupervised Learning for Sentence Compression*
Jenine Turner and Eugene Charniak

10:00–10:30 *Digesting Virtual "Geek" Culture: The Summarization of Technical Internet Relay Chats*
Liang Zhou and Eduard Hovy

10:30–11:00 Break

11:00–12:00 Lifetime Achievement Award and Talk

12:00–1:30 Lunch

1:30–2:30 ACL Business Meeting

Monday, June 27, 2005 (continued)

Session M5R: Parsing

2:30–3:00 *Lexicalization in Crosslinguistic Probabilistic Parsing: The Case of French*
Abhishek Arun and Frank Keller

3:00–3:30 *What to do when Lexicalization Fails: Parsing German with Suffix Analysis and Smoothing*
Amit Dubey

Session M5M: Corpus Annotation

2:30–3:00 *Detecting Errors in Discontinuous Structural Annotation*
Markus Dickinson and W. Detmar Meurers

3:00–3:30 *High Precision Treebanking — Blazing Useful Trees Using POS Information*
Takaaki Tanaka, Francis Bond, Stephan Oepen and Sanae Fujita

3:30–4:00 Break

Session M6R: Machine Learning and Statistical Methods

4:00–4:30 *A Dynamic Bayesian Framework to Model Context and Memory in Edit Distance Learning: An Application to Pronunciation Classification*
Karim Filali and Jeff Bilmes

4:30–5:00 *Learning Stochastic OT Grammars: A Bayesian Approach using Data Augmentation and Gibbs Sampling*
Ying Lin

5:00–5:30 *Contrastive Estimation: Training Log-Linear Models on Unlabeled Data*
Noah A. Smith and Jason Eisner

Monday, June 27, 2005 (continued)

Session M6M: Information Extraction

4:00–4:30 *Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling*
Jenny Rose Finkel, Trond Grenager and Christopher Manning

4:30–5:00 *Unsupervised Learning of Field Segmentation Models for Information Extraction*
Trond Grenager, Dan Klein and Christopher Manning

5:00–5:30 *A Semantic Approach to IE Pattern Induction*
Mark Stevenson and Mark Greenwood

Tuesday, June 28, 2005

Session M7R: Word Sense Disambiguation

9:00–9:30 *Word Sense Disambiguation vs. Statistical Machine Translation*
Marine Carpuat and Dekai Wu

9:30–10:00 *Word Sense Disambiguation Using Label Propagation Based Semi-Supervised Learning*
Zheng-Yu Niu, Dong-Hong Ji and Chew Lim Tan

10:00–10:30 *Domain Kernels for Word Sense Disambiguation*
Alfio Gliozzo, Claudio Giuliano and Carlo Strapparava

Session M7M: Information Extraction

9:00–9:30 *Improving Name Tagging by Reference Resolution and Relation Detection*
Heng Ji and Ralph Grishman

9:30–10:00 *Extracting Relations with Integrated Information Using Kernel Methods*
Shubin Zhao and Ralph Grishman

10:00–10:30 *Exploring Various Knowledge in Relation Extraction*
GuoDong Zhou, Jian Su, Jie Zhang and Min Zhang

Tuesday, June 28, 2005 (continued)

Session M7B: Speech Processing

9:00–9:30 *A Quantitative Analysis of Lexical Differences Between Genders in Telephone Conversations*
Constantinos Boulis and Mari Ostendorf

9:30–10:00 *Position Specific Posterior Lattices for Indexing Speech*
Ciprian Chelba and Alex Acero

10:00–10:30 *Using Conditional Random Fields for Sentence Boundary Detection in Speech*
Yang Liu, Andreas Stolcke, Elizabeth Shriberg and Mary Harper

10:30–11:00 Break

11:00–12:00 Invited Talk by Michael Jordan

12:000–1:30 Lunch

Session M8R: Machine Translation

1:30–2:00 *Log-linear Models for Word Alignment*
Yang Liu, Qun Liu and Shouxun Lin

2:00–2:30 *Alignment Model Adaptation for Domain-Specific Word Alignment*
Hua Wu, Haifeng Wang and Zhanyi Liu

2:30–3:00 *Stochastic Lexicalized Inversion Transduction Grammar for Alignment*
Hao Zhang and Daniel Gildea

Tuesday, June 28, 2005 (continued)

Session M8M: Information Extraction

1:30–2:00 *Multi-Field Information Extraction and Cross-Document Fusion*
Gideon Mann and David Yarowsky

2:00–2:30 *Simple Algorithms for Complex Relation Extraction with Applications to Biomedical IE*
Ryan McDonald, Fernando Pereira, Seth Kulick, Scott Winters, Yang Jin and Pete White

2:30–3:00 *Resume Information Extraction with Cascaded Hybrid Model*
Kun Yu, Gang Guan and Ming Zhou

Session M8B: Speech and Language Modeling

1:30–2:00 *Discriminative Syntactic Language Modeling for Speech Recognition*
Michael Collins, Brian Roark and Murat Saraclar

2:00–2:30 *A Phonotactic Language Model for Spoken Language Identification*
Haizhou Li and Bin Ma

2:30–3:00 *Reading Level Assessment Using Support Vector Machines and Statistical Language Models*
Sarah Schwarm and Mari Ostendorf

3:00–3:30 Break

Session M9R: Machine Translation

3:30–4:00 *Clause Restructuring for Statistical Machine Translation*
Michael Collins, Philipp Koehn and Ivona Kucerova

4:00–4:30 *Machine Translation Using Probabilistic Synchronous Dependency Insertion Grammars*
Yuan Ding and Martha Palmer

4:30–5:00 *Context-Dependent SMT Model using Bilingual Verb-Noun Collocation*
Young-Sook Hwang and Yutaka Sasaki

5:00–5:30 *A Localized Prediction Model for Statistical Machine Translation*
Christoph Tillmann and Tong Zhang

Tuesday, June 28, 2005 (continued)

Session M9M: Segmentation, Tagging, and Semantic Role Labeling

3:30–4:00 *Instance-based Sentence Boundary Determination by Optimization for Natural Language Generation*
Shimei Pan and James Shaw

4:00–4:30 *Arabic Tokenization, Part-of-Speech Tagging and Morphological Disambiguation in One Fell Swoop*
Nizar Habash and Owen Rambow

4:30–5:00 *Semantic Role Labeling Using Different Syntactic Views*
Sameer Pradhan, Wayne Ward, Kadri Hacioglu, James Martin and Daniel Jurafsky

5:00–5:30 *Joint Learning Improves Semantic Role Labeling*
Kristina Toutanova, Aria Haghighi and Christopher Manning

Session M9B: Lexical Acquisition from Corpora

3:30–4:00 *Paraphrasing with Bilingual Parallel Corpora*
Colin Bannard and Chris Callison-Burch

4:00–4:30 *A Nonparametric Method for Extraction of Candidate Phrasal Terms*
Paul Deane

4:30–5:00 *Automatic Acquisition of Adjectival Subcategorization from Corpora*
Jeremy Yallop, Anna Korhonen and Ted Briscoe

5:00–5:30 *Randomized Algorithms and NLP: Using Locality Sensitive Hash Functions for High Speed Noun Clustering*
Deepak Ravichandran, Patrick Pantel and Eduard Hovy

5:30–5:45 Best Paper Award and Closing