ACL 2007

Proceedings of the Interactive Poster and Demonstration Sessions

June 25–27, 2007

Prague, Czech Republic

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

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Preface

The 45th Annual Meeting of the Association for Computational Linguistics, Posters and Demonstrations session was held between the 25th to 27th June 2007 in Prague. This year we had 113 submissions out of which 61 were selected for presentation, resulting in a 54% acceptance rate.

The criteria for acceptance of posters were to describe original work in progress, and to present innovative methodologies used to solve problems in computational linguistics or NLP. 48 posters were accepted.

For demonstrations the criterion for acceptance was the implementation of mature systems or prototypes in which computational linguistics or NLP technologies are used to solve practically important problems. 13 demonstrations were accepted.

I would like to thank the General Conference Chair of ACL 2007, John Carroll, for his insightful suggestions in formulating the call for papers. My gratitude to the members of the Program Committee for their promptness, professionalism and willingness in reviewing more papers than anticipated.

I would like to extend my thanks to the local organisers who accommodated a number of requests speedily making sure that the scheduling and the physical facilities were in place for this event. Last but not least, my special thanks to Scott Piao and Yutaka Sasaki for their help in the preparation of the camera-ready copy of the proceedings.

Sophia Ananiadou
Chair

Organizers

Chair:

Sophia Ananiadou, University of Manchester (UK)

Program Committee:

Timothy Baldwin, University of Melbourne (Australia)
Srinivas Bangalore, AT&, (USA)
Roberto Basili, University of Rome Tor Vergata (Italy)
Walter Daelemans, University of Antwerp (Belgium)
Beatrice Daille, Universite de Nantes (France)
Tomaz Erjavec, Jozef Stefan Institute in Ljubljana (Slovenia)
Katerina Frantzi, University of Aegean (Greece)
Sanda Harabagiu, University of Texas at Dallas (USA)
Jerry Hobbs, USC/ISI (USA)
Alessandro Lenci, Universita di Pisa (Italy)
Evangelos Milios, Dalhousie University (Canada)
Yusuke Miyao, University of Tokyo (Japan)
Kemal Oflazer, Sabanci University (Turkey)
Stelios Piperidis, ILSP (Greece)
Thierry Poibeau, Universite Paris 13 (France)
Paul Rayson, University of Lancaster (UK)
Philip Resnik, University of Maryland (USA)
Fabio Rinaldi, University of Zurich (Switzerland)
Anne de Roeck, Open University (UK)
Frederique Segond, Xerox Research Centre Europe (France)
Kumiko Tanaka-Ishii, University of Tokyo (Japan)
Kentaro Torisawa, JAIST (Japan)
Yoshimasa Tsuruoka, University of Manchester (UK)
Lucy Vanderwende, Microsoft (USA)
Pierre Zweigenbaum, Universite Paris XI (France)

Table of Contents

Program

Demos

Tuesday, June 26

9:00-10:40 Demo Session 1

MIMUS: A Multimodal and Multilingual Dialogue System for the Home Domain
J. Gabriel Amores, Guillermo Pérez and Pilar Manchón

A Translation Aid System with a Stratified Lookup Interface
Takeshi Abekawa and Kyo Kageura

Multimedia Blog Creation System using Dialogue with Intelligent Robot
Akitoshi Okumura, Takahiro Ikeda, Toshihiro Nishizawa, Shin-ichi Ando and Fumihiro Adachi

SemTAG: a platform for specifying Tree Adjoining Grammars and performing TAG-based Semantic Construction
Claire Gardent and Yannick Parmentier

System Demonstration of On-Demand Information Extraction
Satoshi Sekine and Akira Oda

Multilingual Ontological Analysis of European Directives
Gianmaria Ajani, Guido Boella, Leonardo Lesmo, Alessandro Mazzei and Piercarlo Rossi

13:30-15:20 Demo Session 2

zipfR: Word Frequency Modeling in R
Stefan Evert and Marco Baroni

Linguistically Motivated Large-Scale NLP with C&C and Boxer
James Curran, Stephen Clark and Johan Bos

Don't worry about metaphor: affect detection for conversational agents
Catherine Smith, Timothy Rumbell, John Barnden, Robert Hendley, Mark Lee, Alan Wallington and Li Zhang

An efficient algorithm for building a distributional thesaurus (and other Sketch Engine developments)
Pavel Rychlý and Adam Kilgarriff

Tuesday, June 26

Semantic enrichment of journal articles using chemical named entity recognition
Colin R. Batchelor and Peter T. Corbett

An API for Measuring the Relatedness of Words in Wikipedia
Simone Paolo Ponzetto and Michael Strube

Posters

Monday, June 25

15:10-15:45 Poster Session 1

Machine Learning, Corpus and Information Retrieval

Deriving an Ambiguous Words Part-of-Speech Distribution from Unannotated Text
Reinhard Rapp

Support Vector Machines for Query-focused Summarization trained and evaluated on Pyramid data
Maria Fuentes, Enrique Alfonseca and Horacio Rodríguez

A Joint Statistical Model for Simultaneous Word Spacing and Spelling Error Correction for Korean
Hyungjong Noh, Jeong-Won Cha and Gary Geunbae Lee

An Approximate Approach for Training Polynomial Kernel SVMs in Linear Time
Yu-Chieh Wu, Jie-Chi Yang and Yue-Shi Lee

Rethinking Chinese Word Segmentation: Tokenization, Character Classification, or Word-break Identification
Chu-Ren Huang, Petr Šimon, Shu-Kai Hsieh and Laurent Prévot

A Feature Based Approach to Leveraging Context for Classifying Newsgroup Style Discussion Segments
Yi-Chia Wang, Mahesh Joshi and Carolyn Rose

Ensemble document clustering using weighted hypergraph generated by NMF
Hiroyuki Shinnou and Minoru Sasaki

Using Error-Correcting Output Codes with Model-Refinement to Boost Centroid Text Classifier
Songbo Tan

Monday, June 25

Poliqarp: An open source corpus indexer and search engine with syntactic extensions
Daniel Janus and Adam Przepiórkowski

Test Collection Selection and Gold Standard Generation for a Multiply-Annotated Opinion Corpus
Lun-Wei Ku, Yong-Sheng Lo and Hsin-Hsi Chen

Generating Usable Formats for Metadata and Annotations in a Large Meeting Corpus
Andrei Popescu-Belis and Paula Estrella

Exploration of Term Dependence in Sentence Retrieval
Keke Cai, Jiajun Bu, Chun Chen and Kangmiao Liu

Minimum Bayes Risk Decoding for BLEU
Nicola Ehling, Richard Zens and Hermann Ney

10:40-11:10 Poster Session 2

Speech Dialogue

Disambiguating Between Generic and Referential "You" in Dialog
Surabhi Gupta, Matthew Purver and Dan Jurafsky

On the formalization of Invariant Mappings for Metaphor Interpretation
Rodrigo Agerri, John Barnden, Mark Lee and Alan Wallington

Real-Time Correction of Closed-Captions
Patrick Cardinal, Gilles Boulianne, Michel Comeau and Maryse Boisvert

Learning to Rank Definitions to Generate Quizzes for Interactive Information Presentation
Ryuichiro Higashinaka, Kohji Dohsaka and Hideki Isozaki

Predicting Evidence of Understanding by Monitoring User's Task Manipulation in Multimodal Conversations
Yukiko Nakano, Kazuyoshi Murata, Mika Enomoto, Yoshiko Arimoto, Yasuhiro Asa and Hirohiko Sagawa

Automatically Assessing the Post Quality in Online Discussions on Software
Markus Weimer, Iryna Gurevych and Max Mühlhäuser

Monday, June 25

Tuesday, June 26

15:20-15:45 Poster Session 3

Lexica and Ontologies

Tuesday, June 26

An OWL Ontology for HPSG
Graham Wilcock

Wednesday, June 27

10:40-11:10 Poster Session 4

Applications

Classifying Temporal Relations Between Events
Nathanael Chambers, Shan Wang and Dan Jurafsky

Moses: Open Source Toolkit for Statistical Machine Translation
Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondrej Bojar, Alexandra Constantin and Evan Herbst

Boosting Statistical Machine Translation by Lemmatization and Linear Interpolation
Ruiqiang Zhang and Eiichiro Sumita

Extractive Summarization Based on Event Term Clustering
Maofu Liu, Wenjie Li, Mingli Wu and Qin Lu

Machine Translation between Turkic Languages
Ahmet Cüneyd Tantuğ, Esref Adali and Kemal Oflazer

Measuring Importance and Query Relevance in Topic-focused Multi-document Summarization
Surabhi Gupta, Ani Nenkova and Dan Jurafsky

Expanding Indonesian-Japanese Small Translation Dictionary Using a Pivot Language
Masatoshi Tsuchiya, Ayu Purwarianti, Toshiyuki Wakita and Seiichi Nakagawa

Wednesday, June 27

15:10-15:45 Poster Session 5

Parsing and Tagging

Shallow Dependency Labeling
Manfred Klenner

Minimally Lexicalized Dependency Parsing
Daisuke Kawahara and Kiyotaka Uchimoto

HunPos – an open source trigram tagger
Péter Halácsy, András Kornai and Csaba Oravecz

Extending MARIE: an N-gram-based SMT decoder
Josep M. Crego and José B. Mariño

A Hybrid Approach to Word Segmentation and POS Tagging
Tetsuji Nakagawa and Kiyotaka Uchimoto

Automatic Part-of-Speech Tagging for Bengali: An Approach for Morphologically Rich Languages in a Poor Resource Scenario
Sandipan Dandapat, Sudeshna Sarkar and Anupam Basu

Japanese Dependency Parsing Using Sequential Labeling for Semi-spoken Language
Kenji Imamura, Genichiro Kikui and Norihito Yasuda