HLT/EMNLP 2005

Interactive Demonstrations Proceedings

Co-chairs:

Donna Byron, The Ohio State University
Anand Venkataraman, SRI International
Dell Zhang, Birkbeck, University of London

October 7, 2005
Vancouver, British Columbia, Canada

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

The conference organizers are grateful to the following sponsors for their generous support.

Silver Sponsor:

Bronze Sponsors:

Sponsor of Best Student Paper Award:

*This year's HLT/EMNLP conference is co-sponsored by **ACL SIGDAT**.*

Copyright © 2005 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
3 Landmark Center
East Stroudsburg, PA 18301
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Preface to the Demonstration Proceedings

This volume contains the abstracts of the technology demonstrations that were presented at the combined 2005 Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, held in Vancouver, British Columbia, Canada on October 7, 2005. The demonstration program has traditionally been an important component of the HLT conference, and the addition of EMNLP this year creates a new opportunity for interaction among researchers coming from a variety of methodological perspectives. The demonstration program allows researchers the opportunity to showcase new and innovative technology, to update the research community on ongoing projects, and to show software tools developed for corpus development or other research aids to the users they were designed to assist. Demo technologies included in the program were selected based on their potential appeal to the HLT/EMNLP audience, technical and innovative merit, and completeness as a stand-alone demo. Out of 31 demonstration proposals submitted for review, 20 were selected for inclusion at the conference, representing 32 institutions from Asia, North America, and Europe.

We would like to acknowledge Google, Inc. for their generous support of the conference and the demo session. We are also grateful to our reviewers for spending time to help us select a set of demonstrations that together compose a high-quality and informative demo program. We would also like to thank the HLT/EMNLP-2005 conference organizers for their assistance in setting up the program, and Priscilla Rasmussen for local organization.

Donna Byron, Anand Venkataraman and Dell Zhang (editors and co-chairs)
August, 2005

Review Committee:

Harry Bratt, SRI International
Hang Cui, National University of Singapore
Chris Culy, N/A
Marcello Federico, ITC-irst
Eric Fosler-Lussier, The Ohio State University
Andreas Kathol, SRI International
Xiaoli Li, Institute for Infocomm Research, Singapore
Bing Liu, University of Illinois at Chicago
Yang Liu, ICSI, Berkeley
Wen-Hsiang Lu, National Cheng Kung University, Taiwan
Detmar Meurers, The Ohio State University
Murat Saraclar, Bogazici University
Elizabeth Shriberg, SRI International
Gokhan Tur, AT&T Labs
Wen Wang, SRI International
Timothy Weale, The Ohio State University

Table of Contents

Automatic Detection of Translation Errors: The State of the Art
Graham Russell, George Foster and Ngoc Tran Nguyen 1

Bridging the Gap between Technology and Users: Leveraging Machine
Thomas Hoeft, Nick Cramer, M. L. Gregory and Elizabeth Hetzler 2

Classummary: Introducing Discussion Summarization to Online Classrooms
Liang Zhou, Erin Shaw, Chin-Yew Lin and Eduard Hovy 4

Demonstrating an Interactive Semantic Role Labeling System
Vasin Punyakanok, Dan Roth, Mark Sammons and Wen-tau Yih 6

MindNet: An Automatically-Created Lexical Resource
Lucy Vanderwende, Gary Kacmarcik, Hisami Suzuki and Arul Menezes 8

NooJ: a Linguistic Annotation System for Corpus Processing
Max Silberztein 10

Pattern Visualization for Machine Translation Output
Adam Lopez and Philip Resnik 12

Prague Dependency Treebank as an Exercise Book of Czech
Barbora Hladká and Ondřej Kučera 14

Translation Exercise Assistant: Automated Generation of Translation
Jill Burstein and Daniel Marcu 16

WebExperimenter for Multiple-Choice Question Generation
Ayako Hoshino and Hiroshi Nakagawa 18

DialogueView: an Annotation Tool for Dialogue
Fan Yang and Peter A. Heeman 20

Extracting Information about Outbreaks of Infectious Epidemics
Roman Yangarber, Lauri Jokipii, Antti Rauramo and Silja Huttunen 22

A Flexible Conversational Dialog System for MP3 Player
Fuliang Weng, Lawrence Cavedon, Badri Raghunathan, Danilo Mirkovic, Ben Bei, Heather Pon-Barry, Harry Bratt, Hua Cheng, Hauke Schmidt, Rohit Mishra, Brian Lathrop, Qi Zhang, Tobias Scheideck, Kui Xu, Tess Hand-Bender, Stanley Peters, Liz Shriberg and Carsten Bergmann 24

Japanese Speech Understanding using Grammar Specialization
Manny Rayner, Nikos Chatzichrisafis, Pierrette Bouillon, Yukie Nakao, Hitoshi Isahara, Kyoko Kanzaki, Beth Ann Hockey, Marianne Santaholma and Marianne Starlander 26

The MIT Spoken Lecture Processing Project
James R. Glass, Timothy J. Hazen, D. Scott Cyphers, Ken Schutte and Alex Park 28

MBOI: Discovery of Business Opportunities on the Internet
Arman Tajarobi, Jean-François Garneau and François Paradis 30

OPINE: Extracting Product Features and Opinions from Reviews
Ana-Maria Popescu, Bao Nguyen and Oren Etzioni 32

OpinionFinder: A System for Subjectivity Analysis
Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff and Siddharth Patwardhan 34

POSBIOTM/W: A Development Workbench for Machine Learning Oriented Biomedical Text Mining System
Kyungduk Kim, Yu Song and Gary Geunbae Lee 36

Conference Program

Friday, October 7, 2005

Session 1: 6:30-8:00pm

Automatic Detection of Translation Errors: The State of the Art
Graham Russell, George Foster and Ngoc Tran Nguyen

Bridging the Gap between Technology and Users: Leveraging Machine
Thomas Hoeft, Nick Cramer, M. L. Gregory and Elizabeth Hetzler

Classummary: Introducing Discussion Summarization to Online Classrooms
Liang Zhou, Erin Shaw, Chin-Yew Lin and Eduard Hovy

Demonstrating an Interactive Semantic Role Labeling System
Vasin Punyakanok, Dan Roth, Mark Sammons and Wen-tau Yih

MindNet: An Automatically-Created Lexical Resource
Lucy Vanderwende, Gary Kacmarcik, Hisami Suzuki and Arul Menezes

NooJ: a Linguistic Annotation System for Corpus Processing
Max Silberztein

Pattern Visualization for Machine Translation Output
Adam Lopez and Philip Resnik

Prague Dependency Treebank as an Exercise Book of Czech
Barbora Hladká and Ondřej Kučera

Translation Exercise Assistant: Automated Generation of Translation
Jill Burstein and Daniel Marcu

WebExperimenter for Multiple-Choice Question Generation
Ayako Hoshino and Hiroshi Nakagawa

Friday, October 7, 2005 (continued)

Session 2: 8:00-9:30pm

DialogueView: an Annotation Tool for Dialogue
Fan Yang and Peter A. Heeman

Extracting Information about Outbreaks of Infectious Epidemics
Roman Yangarber, Lauri Jokipii, Antti Rauramo and Silja Huttunen

A Flexible Conversational Dialog System for MP3 Player
Fuliang Weng, Lawrence Cavedon, Badri Raghunathan, Danilo Mirkovic, Ben Bei, Heather Pon-Barry, Harry Bratt, Hua Cheng, Hauke Schmidt, Rohit Mishra, Brian Lathrop, Qi Zhang, Tobias Scheideck, Kui Xu, Tess Hand-Bender, Stanley Peters, Liz Shriberg and Carsten Bergmann

Japanese Speech Understanding using Grammar Specialization
Manny Rayner, Nikos Chatzichrisafis, Pierrette Bouillon, Yukie Nakao, Hitoshi Isahara, Kyoko Kanzaki, Beth Ann Hockey, Marianne Santaholma and Marianne Starlander

The MIT Spoken Lecture Processing Project
James R. Glass, Timothy J. Hazen, D. Scott Cyphers, Ken Schutte and Alex Park

MBOI: Discovery of Business Opportunities on the Internet
Arman Tajarobi, Jean-François Garneau and François Paradis

OPINE: Extracting Product Features and Opinions from Reviews
Ana-Maria Popescu, Bao Nguyen and Oren Etzioni

OpinionFinder: A System for Subjectivity Analysis
Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff and Siddharth Patwardhan

POSBIOTM/W: A Development Workbench for Machine Learning Oriented Biomedical Text Mining System
Kyungduk Kim, Yu Song and Gary Geunbae Lee