ACL-05

Interactive Poster and Demonstration Sessions

Proceedings

26 June 2005
University of Michigan

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

©2005 The Association for Computational Linguistics

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

Introduction

The ACL 2005 Interactive Poster and Demonstration session took place on Sunday, June 26, 2005 in Ann Arbor Michigan, on the first day of the 43rd Annual Meeting of the Association for Computational Linguistics. There were 56 submissions to this event, of which 31 were selected for presentation, resulting in a 55% acceptance rate.

Our goal this year was to have a program made up of implemented systems that reflect previously unpublished work (Interactive Posters) or previously published but subsequently improved existing systems (Demonstrations). Thus, while we maintained a fairly traditional definition of *Demonstration*, we introduced an experimental sense of *Interactive Poster* that required that they present novel and previously unpublished ideas *and* that they be supported by an implemented system.

Thus, the criteria for acceptance of an Interactive Poster was the novel scientific contributions of the work, and the effectiveness of the implemented system and (optional) more traditional poster in making those points. There were 37 Interactive Posters submitted, of which 18 were accepted.

Demonstrations were expected to highlight mature systems or prototypes that show how NLP technologies are used to solve practically important problems. The criteria for acceptance of a Demonstration was that it must already be described in the published literature in sufficient detail to allow replication, or the published paper should provide this level of detail. In addition, the Demonstration should address an application of broad interest in such a way that it can be appreciated by the diverse audience that attends ACL. There were 19 Demonstrations submitted, of which 13 were accepted.

First and foremost, we would like to thank the General Conference Chair of ACL 2005, Kevin Knight, for giving us the flexibility and support to pursue a somewhat non–traditional formulation of this event. We would also like to thank the members of the Program Committee for their speedy and insightful reviews, and for their help in refining our definitions of Interactive Poster and Demonstration. Finally, we owe great thanks to Dragomir Radev, who worked tirelessly as the Local Arrangements Chair to accommodate a wide range of requests that we made regarding both the scheduling and physical facilities for this event. We are deeply grateful to him for his hard work which certainly helped to make this a very successful event.

Masaaki Nagata and Ted Pedersen
Co–Chairs

Organizers

Co–Chairs:

Masaaki Nagata (NTT Cyber Space Laboratories)
Ted Pedersen (University of Minnesota, Duluth)

Program Committee:

John Atkinson (Universidad de Concepcion)
Jill Burstein (Educational Testing Service)
Mike Calcagno (Microsoft Speech and Natural Language)
Charles Callaway (University of Edinburgh)
Ciprian Chelba (Microsoft Research)
Silviu Cucerzan (Microsoft Research)
Mark Dras (Macquarie University)
Gregor Erbach (DFKI GmbH)
Reva Freedman (Northern Illinois University)
Jianfeng Gao (Microsoft Research Asia)
Alexander Gelbukh (National Polytechnic Institute, Mexico)
Alexander Hauptmann (Carnegie Mellon University)
Yoshihiko Hayashi (Osaka University)
Graeme Hirst (University of Toronto)
Camelia Ignat (European Commission - Joint Research Centre)
Diana Inkpen (University of Ottawa)
Su Jian (Institute for Infocomm Research)
Jin-Dong Kim (University of Tokyo)
Taku Kudo (Google)
Sadao Kurohashi (University of Tokyo)
Claudia Leacock (Pearson Knowledge Technologies)
Gary Geunbae Lee (Pohang University of Science and Technology)
Diane Litman (University of Pittsburgh)
Sun Maosong (Tsinghua University)
Rada Mihalcea (University of North Texas)
Vibhu Mittal (Google/Carnegie Mellon University)
Mikio Nakano (Honda Research Institute Japan)
Manabu Okumura (Tokyo Institute of Technology)
Serguei Pakhomov (Mayo College of Medicine)
Patrick Pantel (USC Information Sciences Institute)
Marius Pasca (Google)
David Pierce (University at Buffalo)
Bruno Pouliquen (European Commission - Joint Research Centre)
Philip Resnik (University of Maryland)
Carolyn Penstein Rosé (Carnegie Mellon University)
Horacio Saggion (University of Sheffield)

William Schuler (University of Minnesota, Twin Cities)
Kiyoaki Shirai (Japan Advanced Institute of Science and Technology)
Thamar Solorio (Instituto Nacional de Astrofísica, Óptica, y Electrónica)
Virach Sornlertlamvanich (NICT Asia Research Center)
Ralf Steinberger (European Commission - Joint Research Centre)
Keh-Yih Su (Behavior Design Corporation)
Peter Turney (National Research Council of Canada)
Masao Utiyama (NICT)
Takehito Utsuro (Kyoto University)
Janyce Wiebe (University of Pittsburgh)
Peter Wiemer-Hastings (DePaul University)

Additional Reviewers:

John Chen (Microsoft Research Asia)
Hal Daume III (USC Information Sciences Institute)
Alexander Fraser (USC Information Sciences Institute)
Zhang Jie (Institute for Infocomm Research)
Mu Li (Microsoft Research Asia)
Yajuan Lv (Microsoft Research Asia)
Liang Zhou (USC Information Sciences Institute)

Table of Contents

An Information-State Approach to Collaborative Reference
David DeVault, Natalia Kariaeva, Anubha Kothari, Iris Oved and Matthew Stone 1

Accessing GermaNet Data and Computing Semantic Relatedness
Iryna Gurevych and Hendrik Niederlich 5

Efficient Solving and Exploration of Scope Ambiguities
Alexander Koller and Stefan Thater 9

CL Research's Knowledge Management System
Kenneth C. Litkowski 13

Dynamically Generating a Protein Entity Dictionary Using Online Resources
Hongfang Liu, Zhangzhi Hu and Cathy Wu 17

Descriptive Question Answering in Encyclopedia
Hyo-Jung O. Lee, Hyeon-Jin Kim and Myung-Gil Jang 21

High Throughput Modularized NLP System for Clinical Text
Serguei Pakhomov, James Buntrock and Patrick Duffy 25

A Voice Enabled Procedure Browser for the International Space Station
Manny Rayner, Beth A. Hockey, Nikos Chatzichrisafis, Kim Farrell and Jean-Michel Renders . 29

The Linguist's Search Engine: An Overview
Philip Resnik and Aaron Elkiss 33

Learning Source-Target Surface Patterns for Web-based Terminology Translation
Jian-Cheng Wu, Tracy Lin and Jason S. Chang 37

SPEECH OGLE: Indexing Uncertainty for Spoken Document Search
Ciprian Chelba and Alex Acero 41

Multimodal Generation in the COMIC Dialogue System
Mary E. Foster, Michael White, Andrea Setzer and Roberta Catizone 45

Language Independent Extractive Summarization
Rada Mihalcea 49

SenseLearner: Word Sense Disambiguation for All Words in Unrestricted Text
Rada Mihalcea and Andras Csomai 53

Syntax-based Semi-Supervised Named Entity Tagging
Behrang Mohit and Rebecca Hwa 57

Portable Translator Capable of Recognizing Characters on Signboard and Menu Captured by its Built-in Camera
Hideharu Nakajima, Yoshihiro Matsuo, Masaaki Nagata and Kuniko Saito 61

Supporting Annotation Layers for Natural Language Processing
Preslav Nakov, Ariel Schwartz, Brian Wolf and Marti Hearst 65

Word Alignment and Cross-Lingual Resource Acquisition
Carol Nichols and Rebecca Hwa 69

SenseRelate::TargetWord - A Generalized Framework for Word Sense Disambiguation
Siddharth Patwardhan, Satanjeev Banerjee and Ted Pedersen 73

A Practical Solution to the Problem of Automatic Part-of-Speech Induction from Text
Reinhard Rapp 77

Automating Temporal Annotation with TARSQI
Marc Verhagen, Inderjeet Mani, Roser Sauri, Jessica Littman, Robert Knippen, Seok B. Jang, Anna Rumshisky, John Phillips and James Pustejovsky 81

Two Diverse Systems Built using Generic Components for Spoken Dialogue (Recent Progress on TRIPS)
James Allen, George Ferguson, Amanda Stent, Scott Stoness, Mary Swift, Lucian Galescu, Nathan Chambers, Ellen Campana and Gregory Aist 85

Transonics: A Practical Speech-to-Speech Translator for English-Farsi Medical Dialogs
Robert Belvin, Emil Ettelaie, Sudeep Gandhe, Panayiotis Georgiou, Kevin Knight, Daniel Marcu, Scott Millward, Shrikanth Narayanan, Howard Neely and David Traum 89

The Wild Thing
Ken Church and Bo Thiesson 93

Interactively Exploring a Machine Translation Model
Steve DeNeefe, Kevin Knight and Hayward H. Chan 97

Multi-Engine Machine Translation Guided by Explicit Word Matching
Shyamsundar Jayaraman and Alon Lavie 101

SenseClusters: Unsupervised Clustering and Labeling of Similar Contexts
Anagha Kulkarni and Ted Pedersen 105

A Flexible Stand-Off Data Model with Query Language for Multi-Level Annotation
Christoph Mueller 109

HAHAcronym: A Computational Humor System
Oliviero Stock and Carlo Strapparava 113

Organizing English Reading Materials for Vocabulary Learning
Masao Utiyama, Midori Tanimura and Hitoshi Isahara 117

Reformatting Web Documents via Header Trees
Minoru Yoshida and Hiroshi Nakagawa 121

Interactive Poster (IP) and Demonstration (D) Program

Sunday, June 26, 2005 (10:00 am – 12:00 noon)

An Information-State Approach to Collaborative Reference (IP)
David DeVault, Natalia Kariaeva, Anubha Kothari, Iris Oved and Matthew Stone

Accessing GermaNet Data and Computing Semantic Relatedness (IP)
Iryna Gurevych and Hendrik Niederlich

Efficient Solving and Exploration of Scope Ambiguities (D)
Alexander Koller and Stefan Thater

CL Research's Knowledge Management System (IP)
Kenneth C. Litkowski

Dynamically Generating a Protein Entity Dictionary Using Online Resources (D)
Hongfang Liu, Zhangzhi Hu and Cathy Wu

Descriptive Question Answering in Encyclopedia (IP)
Hyo-Jung O. Lee, Hyeon-Jin Kim and Myung-Gil Jang

High Throughput Modularized NLP System for Clinical Text (IP)
Serguei Pakhomov, James Buntrock and Patrick Duffy

A Voice Enabled Procedure Browser for the International Space Station (D)
Manny Rayner, Beth A. Hockey, Nikos Chatzichrisafis, Kim Farrell and Jean-Michel Renders

The Linguist's Search Engine: An Overview (D)
Philip Resnik and Aaron Elkiss

Learning Source-Target Surface Patterns for Web-based Terminology Translation (IP)
Jian-Cheng Wu, Tracy Lin and Jason S. Chang

Sunday, June 26, 2005 (1:30 pm – 3:30 pm)

SPEECH OGLE: Indexing Uncertainty for Spoken Document Search (IP)
Ciprian Chelba and Alex Acero

Multimodal Generation in the COMIC Dialogue System (D)
Mary E. Foster, Michael White, Andrea Setzer and Roberta Catizone

Language Independent Extractive Summarization (D)
Rada Mihalcea

SenseLearner: Word Sense Disambiguation for All Words in Unrestricted Text (IP)
Rada Mihalcea and Andras Csomai

Syntax-based Semi-Supervised Named Entity Tagging (IP)
Behrang Mohit and Rebecca Hwa

Portable Translator Capable of Recognizing Characters on Signboard and Menu Captured by its Built-in Camera (D)
Hideharu Nakajima, Yoshihiro Matsuo, Masaaki Nagata and Kuniko Saito

Supporting Annotation Layers for Natural Language Processing (IP)
Preslav Nakov, Ariel Schwartz, Brian Wolf and Marti Hearst

Word Alignment and Cross-Lingual Resource Acquisition (IP)
Carol Nichols and Rebecca Hwa

SenseRelate::TargetWord - A Generalized Framework for Word Sense Disambiguation (D)
Siddharth Patwardhan, Satanjeev Banerjee and Ted Pedersen

A Practical Solution to the Problem of Automatic Part-of-Speech Induction from Text (IP)
Reinhard Rapp

Automating Temporal Annotation with TARSQI (IP)
Marc Verhagen, Inderjeet Mani, Roser Sauri, Jessica Littman, Robert Knippen, Seok B. Jang, Anna Rumshisky, John Phillips and James Pustejovsky

Sunday, June 26, 2005 (4:00 pm – 6:00 pm)

Two Diverse Systems Built using Generic Components for Spoken Dialogue (Recent Progress on TRIPS) (IP)
James Allen, George Ferguson, Amanda Stent, Scott Stoness, Mary Swift, Lucian Galescu, Nathan Chambers, Ellen Campana and Gregory Aist

Transonics: A Practical Speech-to-Speech Translator for English-Farsi Medical Dialogs (D)
Robert Belvin, Emil Ettelaie, Sudeep Gandhe, Panayiotis Georgiou, Kevin Knight, Daniel Marcu, Scott Millward, Shrikanth Narayanan, Howard Neely and David Traum

The Wild Thing (D)
Ken Church and Bo Thiesson

Interactively Exploring a Machine Translation Model (IP)
Steve DeNeefe, Kevin Knight and Hayward H. Chan

Multi-Engine Machine Translation Guided by Explicit Word Matching (IP)
Shyamsundar Jayaraman and Alon Lavie

SenseClusters: Unsupervised Clustering and Labeling of Similar Contexts (D)
Anagha Kulkarni and Ted Pedersen

A Flexible Stand-Off Data Model with Query Language for Multi-Level Annotation (IP)
Christoph Mueller

HAHAcronym: A Computational Humor System (D)
Oliviero Stock and Carlo Strapparava

Organizing English Reading Materials for Vocabulary Learning (D)
Masao Utiyama, Midori Tanimura and Hitoshi Isahara

Reformatting Web Documents via Header Trees (IP)
Minoru Yoshida and Hiroshi Nakagawa