

**Spoken Language Understanding
for Conversational Systems
and
Higher Level Linguistic Information
for Speech Processing**

Proceedings of the HLT-NAACL 2004 Workshop

**7 May 2004
Boston, USA**

**Spoken Language Understanding
for Conversational Systems
and
Higher Level Linguistic Information
for Speech Processing**

Proceedings of the HLT-NAACL 2004 Workshop

**Co-chairs:
Srinivas Bangalore
Yuqing Gao
Hong-Kwang Jeff Kuo
Andreas Stolcke
Dilek Hakkani-Tür
Gokhan Tur**

**7 May 2004
Boston, USA**

©2004 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

The Association for Computational Linguistics
3 Landmark Center
East Stroudsburg, PA 18301
USA
acl@aclweb.org

Preface

This volume contains the papers accepted for presentation at the workshop on *Spoken Language Understanding in Conversational Systems and Higher Level Linguistic Knowledge for Speech Processing* organized at the Human Language Technology conference/North American chapter of the Association for Computational Linguistics annual meeting (HLT/NAACL 2004).

The success of a conversational system depends on a synergistic integration of technologies such as speech recognition, spoken language understanding (SLU), dialog modeling, natural language generation, speech synthesis and user interface design. In this workshop, we address the issue of improving the *robustness* of the speech recognition and SLU components by exploiting higher level linguistic knowledge, meta-information and machine learning techniques.

The challenging robustness issues in speech recognition such as compensation for acoustic confusability resulting from noisy environments and unexpected channel and speaker mismatch can potentially be aided by the use of higher-level linguistic information such as prosody, syntax, semantics, and pragmatics and even high-level meta-information, such as personal information stored in a database or dialogue and pragmatic coherence constraints. However, current state-of-the-art speech recognizers do not explicitly use such information and rely mainly on information encoded in statistical N-gram language models. Furthermore, high-level information is typically crucial when the ultimate goal is to interpret the spoken input (i.e., the same sequence of words can mean different things depending on prosodic and syntactic features, as well as pragmatic constraints).

The topic of robust SLU has received much attention during the DARPA funded ATIS program of the 1990s and more recently the DARPA Communicator program. In parallel to that research, a number of real-world conversational systems have been deployed to date. However, the techniques for robust SLU have branched out in many different directions. They have been influenced by many recent areas such as information extraction, question answering and machine learning. Data driven approaches to understanding are rapidly gaining prominence. There has been substantial increase in interest in information extraction from the NLP community, question-answering in the information retrieval community, and spoken dialog systems in the speech processing community. Spoken language understanding is an especially attractive topic for cross-fertilization of ideas between Speech, IR and NLP communities.

We thank Prof. Renato de Mori, University of Avignon and Dr. Roberto Pieraccini, IBM, for accepting our invitation to share their ideas on spoken language processing and conversational systems. We thank the program committee members for their informative reviews of the submitted papers. We thank the authors for electing to present their work at this forum. We finally thank the HLT/NAACL for supporting this workshop.

INVITED SPEAKERS:

Renato de Mori, University of Avignon, France
Roberto Pieraccini, IBM TJ Watson Research Center, USA

ORGANIZERS:

Srinivas Bangalore, AT&T Labs - Research
Dilek Hakkani-Tür, AT&T Labs - Research
Gokhan Tur, AT&T Labs - Research
Yuqing Gao, IBM TJ Watson Research Center
Hong-Kwang Jeff Kuo, IBM TJ Watson Research Center
Andreas Stolcke, SRI & ICSI

PROGRAM COMMITTEE:

Frederic Bechet, Univ. of Avignon, France
Jerome Bellegarda, Apple Computer, USA
Jennifer Chu-Carroll, IBM TJ Watson Research Center, USA
Ciprian Chelba, Microsoft, USA
Stephen Cox, Univ. of East Anglia, UK
Sadaoki Furui, Tokyo Institute of Technology, Japan
Allen Gorin, AT&T Labs - Research, USA
Roberto Gretter, ITC-IRST, Italy
Julia Hirschberg, Columbia University, USA
Dan Jurafsky, University of Colorado, USA
Sanjeev Khudanpur, Johns Hopkins University, USA
Helen Meng, CUHK, Hong Kong
Prem Natarajan, BBN, USA
Hermann Ney, RWTH Aachen, Germany
Martha Palmer, University of Pennsylvania, USA
Barbara Peskin, ICSI, USA
Roberto Pieraccini, IBM TJ Watson Research Center, USA
Manny Rayner, NASA, USA
Brian Roark, AT&T Labs - Research, USA
Roni Rosenfeld, Carnegie Mellon University, USA
Stephanie Seneff, MIT, USA
Elizabeth Shriberg, SRI, USA
Amanda Stent, Stony Brook Univ., USA

CONFERENCE WEBSITES:

<http://www.research.att.com/~dtur/NAACL04-Workshop/>
<http://www.speech.sri.com/hlt-workshop/>

CONFERENCE PROGRAM

Friday, May 7

- 8:45-9:00 Welcome
- 9:00-9:50 *Invited Talk: Sentence Interpretation using Stochastic Finite State Transducers*
Renato De Mori
- 9:50-10:00 Break
- 10:00-10:30 *Hybrid Statistical and Structural Semantic Modeling for Thai Multi-Stage Spoken Language Understanding*
Chai Wutiw WATCHAI and Sadaoki FURUI
- 10:30-11:00 *Interactive Machine Learning Techniques for Improving SLU Models*
Lee Begeja, Bernard Renger, David Gibbon, Zhu Liu and Behzad Shahraray
- 11:00-11:30 *Virtual Modality: a Framework for Testing and Building Multimodal Applications*
Péter Pál Boda and Edward Filisko
- 11:30-12:00 *Automatic Call Routing with Multiple Language Models*
Qiang Huang and Stephen Cox
- 12:00-1:00 Lunch
- 1:00-1:30 *Error Detection and Recovery in Spoken Dialogue Systems*
Edward Filisko and Stephanie Seneff
- 1:30-2:00 *Robustness Issues in a Data-Driven Spoken Language Understanding System*
Yulan He and Steve Young
- 2:00-2:50 *Invited Talk: Spoken Language Understanding: the Research/Industry Chasm*
Roberto Pieraccini
- 2:50-3:00 Break
- 3:00-3:30 *Using Higher-level Linguistic Knowledge for Speech Recognition Error Correction in a Spoken Q/A Dialog*
Minwoo Jeong, Byeongchang Kim and Gary Geunbae Lee
- 3:30-4:00 *Speech Recognition Models of the Interdependence Among Syntax, Prosody, and Segmental Acoustics*
Mark Hasegawa-Johnson, Jennifer Cole, Chilin Shih, Ken Chen, Aaron Cohen, Sandra Chavarria, Heejin Kim, Taejin Yoon, Sarah Borys and Jeung-Yoon Choi
- 4:00-4:30 *Modeling Prosodic Consistency for Automatic Speech Recognition: Preliminary Investigations*
Ernest Pusateri and James Glass
- 4:30-5:00 *Assigning Domains to Speech Recognition Hypotheses*
Klaus Rüggenmann and Iryna Gurevych
- 5:00-5:30 *Context Sensing using Speech and Common Sense*
Nathan Eagle and Push Singh

Author Index

Begeja, Lee	10
Boda, Péter Pál	17
Borys, Sarah	56
Chavarria, Sandra	56
Chen, Ken	56
Choi, Jeung-Yoon	56
Cohen, Aaron	56
Cole, Jennifer	56
Cox, Stephen	25
De Mori, Renato	1
Eagle, Nathan	78
Filisko, Edward	17, 31
Furui, Sadaoki	2
Gibbon, David	10
Glass, James	64
Gurevych, Iryna	70
Hasegawa-Johnson, Mark	56
He, Yulan	39
Huang, Qiang	25
Jeong, Minwoo	48
Kim, Byeongchang	48
Kim, Heejin	56
Lee, Gary Geunbae	48
Liu, Zhu	10
Pieraccini, Roberto	47
Pusateri, Ernest	64
Renger, Bernard	10
Rüggenmann, Klaus	70
Seneff, Stephanie	31
Shahraray, Behzad	10
Shih, Chilin	56
Singh, Push	78
Wutiwiwatchai, Chai	2
Yoon, Taejin	56
Young, Steve	39