ACL-08: HLT

Software Engineering, Testing, and Quality Assurance for Natural Language Processing

June 20, 2008
The Ohio State University
Columbus, Ohio, USA

Production and Manufacturing by
Omnipress Inc.
2600 Anderson Street
Madison, WI 53707
USA

©2008 The Association for Computational Linguistics

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

ISBN 978-1-932432-10-7

Software engineering, testing, and quality assurance for natural language processing

Software engineering in general is a first-class research object in computer science, but generally has not been treated as such within the natural language processing community. This is despite the fact that natural language as an input type has unique characteristics that present special problems for software testing, quality assurance, and even requirements specification.

The goals of this workshop included raising awareness of the need for good software engineering practices in NLP, stimulating research on same, and disseminating the results of current work in this area. We are grateful to the authors for sharing their work, and to the program committee for their efforts.

Kevin Bretonnel Cohen and Bob Carpenter

Organizers:

K. Bretonnel Cohen, The MITRE Corporation and University of Colorado School of Medicine
Bob Carpenter, Alias-i

Program Committee:

William A. Baumgartner, Jr., University of Colorado School of Medicine
Hamish Cunningham, University of Sheffield
Dan Flickinger, Stanford University
Michael Gamon, Microsoft
Martin Jansche, Google
Marc Light, Thomson
James Lyle, Microsoft
Kevin Markey, Silver Creek Systems
Stephan Oepen, Stanford University
Martha Palmer, University of Colorado at Boulder
Jeff Reynar, Google
Jun'ichi Tsujii, University of Tokyo and UK National Centre for Text Mining
Martin Volk, University of Stockholm
Scott A. Waterman, PowerSet

Additional Reviewers:

Three additional anonymous reviewers provided reviews of a submission from a workshop organizer.

Table of Contents

Increasing Maintainability of NLP Evaluation Modules Through Declarative Implementations
Terry Heinze and Marc Light 1

Type-checking in Formally Non-typed Systems
Dick Crouch and Tracy Holloway King 3

zymake: A Computational Workflow System for Machine Learning and Natural Language Processing
Eric Breck 5

Evaluating the Effects of Treebank Size in a Practical Application for Parsing
Kenji Sagae, Yusuke Miyao, Rune Saetre and Jun'ichi Tsujii 14

Adapting Naturally Occurring Test Suites for Evaluation of Clinical Question Answering
Dina Demner-Fushman 21

Software Testing and the Naturally Occurring Data Assumption in Natural Language Processing
K. Bretonnel Cohen, William A. Baumgartner Jr. and Lawrence Hunter 23

Building a BioWordNet Using WordNet Data Structures and WordNet's Software Infrastructure–A Failure Story
Michael Poprat, Elena Beisswanger and Udo Hahn 31

Fast, Scalable and Reliable Generation of Controlled Natural Language
David Hardcastle and Richard Power 40

Parallel Implementations of Word Alignment Tool
Qin Gao and Stephan Vogel 49

Design of the Moses Decoder for Statistical Machine Translation
Hieu Hoang and Philipp Koehn 58

Buckwalter-based Lookup Tool as Language Resource for Arabic Language Learners
Jeffrey Micher and Clare Voss 66

Reengineering a Domain-Independent Framework for Spoken Dialogue Systems
Filipe Martins, Ana Mendes, Márcio Viveiros, Joana Paulo Pardal, Pedro Arez, Nuno J. Mamede, and João Paulo Neto 68

Workshop Program

Friday, June 20, 2008

9:00–9:10 *Welcome and opening remarks*

9:10–9:30 *Increasing Maintainability of NLP Evaluation Modules Through Declarative Implementations*
Terry Heinze and Marc Light

9:30–9:50 *Type-checking in Formally Non-typed Systems*
Dick Crouch and Tracy Holloway King

9:50–10:10 *zymake: A Computational Workflow System for Machine Learning and Natural Language Processing*
Eric Breck

10:10–10:30 *Evaluating the Effects of Treebank Size in a Practical Application for Parsing*
Kenji Sagae, Yusuke Miyao, Rune Saetre and Jun'ichi Tsujii

10:30–11:00 **Coffee break**

11:00–11:20 *Adapting Naturally Occurring Test Suites for Evaluation of Clinical Question Answering*
Dina Demner-Fushman

11:20–11:40 *Software Testing and the Naturally Occurring Data Assumption in Natural Language Processing*
K. Bretonnel Cohen, William A. Baumgartner Jr. and Lawrence Hunter

11:40–12:00 *Building a BioWordNet Using WordNet Data Structures and WordNet's Software Infrastructure–A Failure Story*
Michael Poprat, Elena Beisswanger and Udo Hahn

12:00–2:00 **Lunch**

Friday, June 20, 2008 (continued)

2:00–2:20 *Fast, Scalable and Reliable Generation of Controlled Natural Language*
David Hardcastle and Richard Power

2:20–2:40 *Parallel Implementations of Word Alignment Tool*
Qin Gao and Stephan Vogel

2:40–3:00 *Design of the Moses Decoder for Statistical Machine Translation*
Hieu Hoang and Philipp Koehn

3:00–3:20 *Buckwalter-based Lookup Tool as Language Resource for Arabic Language Learners*
Jeffrey Micher and Clare Voss

3:30–4:00 **Coffee break**

4:00–4:20 *Reengineering a Domain-Independent Framework for Spoken Dialogue Systems*
Filipe M. Martins, Ana Mendes, Márcio Viveiros, Joana Paulo Pardal, Pedro Arez, Nuno J. Mamede and João Paulo Neto

4:20–5:00 **Open discussion session**