ACL-05

Workshop on Software

Proceedings of the Workshop

30 June 2005
University of Michigan
Ann Arbor, Michigan, USA

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

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

Welcome to the ACL Workshop on Software, the first of its kind. It is intended as a venue for discussing and comparing the implementation of software and algorithms used in Natural Language and Speech Processing. The goal is to bring together researchers, software developers, teachers, and students with a common interest in the implementation of NLP applications, and to allow useful implementation techniques and "tricks of the trade" to be discussed in detail and disseminated widely.

We received 13 submissions, of which 8 were selected for presentation and inclusion in the proceedings, after a careful review process. Because the number of reviewers exceeded the number of submissions, each submission received more than four reviews on average, while the workload per reviewer was less than three papers on average. This being the workshop on software, the initial assignment of reviews was performed algorithmically using the Ford–Fulkerson max-flow algorithm, while taking into account individual reviewer preferences. The reviewers did an admirable job dealing with a diverse set of submissions, for which they deserve the thanks of the community.

The papers presented in this workshop deal with many different aspects of NLP software: Carpenter describes a scalable implementation of high-order character language models; Clegg & Shepherd take three existing parsers that were trained on business news text and perform a comparative evaluation on a corpus of biomedical journal papers; Cohen-Sygal & Wintner have implemented a compiler which translates between the description languages of two different finite state toolboxes; Foster has designed a generation module for a dialogue system which can ship out text without having to wait for the planning phase to finish; Koller & Thater describe the intelligent design of increasingly powerful constraint solvers; Newman proposes a uniform formalism for representing the output of parsers for easy inspection and comparison; Trón, Gyepesi, Halácsy, Kornai, Németh & Varga have implemented a generic library for analyzing orthographic words; and White discusses the design of a generation component which flexibly incorporates language models in a syntactic surface realizer.

The workshop proceedings are being made available in electronic form only. Not only does this save costs, but it also allows the distribution of additional software and resources that could not be included in printed proceedings. A number of authors have included the software described in their papers directly on the proceedings CD. As always, the latest versions of the included software can be found on the Internet or by contacting the individual authors.

I would like to thank the reviewers and authors once again for their hard work and look forward to an exciting workshop.

Martin Jansche
Columbia University, New York

Organizer:

Martin Jansche, Columbia University (USA)

Program Committee:

Cyril Allauzen, AT&T (USA)
Jason Baldridge, University of Edinburgh (UK)
Srinivas Bangalore, AT&T (USA)
Frédéric Bechet, Université d'Avignon (France)
Tilman Becker, DFKI (Germany)
Steven Bird, University of Melbourne (Australia)
Antal van den Bosch, Universiteit van Tilburg (Netherlands)
Bob Carpenter, Alias-i (USA)
Nizar Habash, Columbia University (USA)
Benoit Lavoie, CoGenTex and Université du Québec à Montréal (Canada)
Alexis Nasr, Université Paris 7 (France)
Hermann Ney, RWTH Aachen (Germany)
Stephan Oepen, CSLI (USA)
Owen Rambow, Columbia University (USA)
Brian Roark, OGI/OHSU (USA)
Richard Sproat, University of Illinois (USA)
Nathan Vaillette, Universität Tübingen (Germany)
Michael White, University of Edinburgh (UK)

Additional Reviewers:

Oliver Bender, RWTH Aachen (Germany)
Evgeny Matusov, RWTH Aachen (Germany)
David Vilar, RWTH Aachen (Germany)

Table of Contents