

CALL FOR PAPERS

ANNOTATION SCIENCE: STATE OF THE ART IN ENHANCING AUTOMATIC LINGUISTIC ANNOTATION

To be held in conjunction with:

The 5th International Language Resources and Evaluation Conference

Magazzini del Cotone Conference Centre

Genoa, Italy

May 27, 2006

The high cost of manual annotation and validation of automatically produced annotations for language data has led to the recent development of methods to enhance the quality of automatically-produced annotations via mechanisms such as machine learning. To date, there has been no international forum fully dedicated to the topic, where researchers working in different areas and different phenomena are brought together to discuss methods and results.

This workshop will include papers describing current work on enhancing the results of automated annotation for linguistic phenomena and provide both an overview and assessment of the state-of-the-art. The workshop will include an introductory overview as well as a panel discussion following the paper presentations.

Papers are solicited on any of the following topics:

- Machine learning and other methods to enhance automatic annotation of linguistic phenomena, including segmentation and chunking; morpho-syntactic, syntactic, and discourse analysis; semantic annotation; entity and event recognition; alignment of parallel translations; annotation of dialogue, speech, gesture, and multi-modal data; etc.
- Use of information from multiple linguistic levels and/or phenomena to enhance performance of automatic annotation software
- Machine learning and other methods for enhancing automated knowledge acquisition (e.g, information for lexicons, ontologies, etc.)
- Evaluation and comparison of techniques to enhance the accuracy of automatically-generated annotations, as well as discussion of limitations
- Software systems for optimizing annotation accuracy, and methods and systems for optimizing "on the fly" annotation of web and other language data
- Identification and separation of annotation that cannot be automated, to simplify annotation enhancement by human proofreaders
- Identification and separation of applications that can tolerate "noisy" analysis, for which imperfect automated linguistic analysis would be appropriate

DATES

Paper submission : March 8, 2006

Author notification : April 7, 2006

Workshop date : May 27, 2006

SUBMISSION INFORMATION

Papers should be no more than 8 pages in length and follow the format for submissions to the main LREC conference. Papers should provide a thorough account of methodology and evaluation of results. Submissions in pdf format should be sent to annotation-science@cs.vassar.edu.

CONTACT

Please send inquiries to annotation-science@cs.vassar.edu.

ORGANIZERS

Nancy Ide
Department of Computer Science
Vassar College
Poughkeepsie, New York 12604-0520 USA
tel: (+1 845) 437 5988
fax: (+1 845) 437 7498
email: ide@cs.vassar.edu

Eric Atwell
School of Computing
University of Leeds
United Kingdom
tel: +44 (0)113 3435761
fax: +44 (0)113 3435468
email: eric@comp.leeds.ac.uk

PROGRAM COMMITTEE

Eric Atwell, University of Leeds (UK)
Nigel Collier, National Institute of Informatics (Japan)
Atsushi Fujii, University of Tsukuba (Japan)
Rebecca Hwa, University of Pittsburgh (USA)
Nancy Ide, Vassar College (USA)
David Lewis, David D. Lewis Consulting USA)
Miles Osborne, University of Edinburgh (UK)
Anoop Sarkar, Simon Fraser University (CA)
Mark Steedman, University of Edinburgh (UK)
Takenobu Tokunaga, Tokyo Institute of Technology (Japan)
Kiyotaka Uchimoto, National Institute of Information and Communications Technology (Japan)