NAACL HLT 2009

BioNLP 2009

Companion Volume:
Shared Task on Event Extraction

June 5, 2009
Boulder, Colorado

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

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 978-1-932432-44-2

Introduction

The need for automatic processing of the rapidly increasing body of publications describing results in molecular biology continues to drive efforts in Biomedical natural language processing (BioNLP). Until recently, domain efforts have largely concentrated on foundational tasks such as entity recognition and relatively simple information extraction targets such as interacting entity pairs. By contrast, biological research increasingly aims to create detailed descriptions of complex processes in biological systems. To respond to the needs of such research, it is necessary to develop BioNLP methods that are able to process more fine-grained representations. The BioNLP'09 Shared Task is the first community-wide step in this direction.

Shared tasks have a strong tradition in the BioNLP community. The TREC Genomics, KDD cup, JNLBPA, LLL and BioCreative tasks have focused the efforts of the community on timely challenges in the domain, both establishing the capabilities and problem points of current systems as well as advancing the state of the art in various areas of biomedical text mining. These are also the goals of the present shared task. The focus on a rich representation of extracted information in the BioNLP'09 Shared Task can also be seen as natural continuation of the succession of previous tasks.

The BioNLP'09 Shared Task targets event extraction following a model similar to those currently applied in the wider IE community. Corpus resources supporting this type of representation have only recently become available in the domain, and the task thus represents a novel challenge to much of the community. Despite its novelty and relatively complex task settings, the BioNLP'09 Shared Task met with an enthusiastic response from the community: more than 40 teams registered their preliminary interest, and 24 teams submitted final results. Thanks to the efforts of the participants and the shared task program committee in reviewing, we have the pleasure of presenting these proceedings of 19 manuscripts accepted for presentation at the shared task session of the workshop.

Shared Task Chair:

Jun'ichi Tsujii, University of Tokyo, University of Manchester, and National Centre for Text Mining

Organizers:

Jin-Dong Kim, Univerisity of Tokyo
Tomoko Ohta, Univerisity of Tokyo
Sampo Pyysalo, Univerisity of Tokyo
Yoshinobu Kano, Univerisity of Tokyo

Program Committee:

Andrew B. Clegg, University College London
Nigel Collier, National Institute of Informatics
Barry Haddow, University of Edinburgh
Martin Krallinger, CNIO
Ryan McDonald, Google
Conrad Plake, Biotechnology Center of TU Dresden
Thierry Poibeau, LIPN-CNRS
Burr Settles, University of Wisconsin-Madison
Yoshimasa Tsuruoka, University of Manchester
Hong Yu, University of Wisconsin-Milwaukee

Table of Contents

Conference Program

Friday, June 5, 2009

Session 1: Oral presentations

9:00–9:20 *Overview of BioNLP'09 Shared Task on Event Extraction*
Jin-Dong Kim, Tomoko Ohta, Sampo Pyysalo, Yoshinobu Kano and Jun'ichi Tsujii

9:20–9:40 *Extracting Complex Biological Events with Rich Graph-Based Feature Sets*
Jari Björne, Juho Heimonen, Filip Ginter, Antti Airola, Tapio Pahikkala and Tapio Salakoski

9:40–10:00 *Event Extraction from Trimmed Dependency Graphs*
Ekaterina Buyko, Erik Faessler, Joachim Wermter and Udo Hahn

10:00–10:15 *UZurich in the BioNLP 2009 Shared Task*
Kaarel Kaljurand, Gerold Schneider and Fabio Rinaldi

10:15–10:30 *Biomedical Event Extraction without Training Data*
Andreas Vlachos, Paula Buttery, Diarmuid Ó Séaghdha and Ted Briscoe

10:30–11:00 Coffee break

Session 2: Oral presentations

11:00–11:20 *A Markov Logic Approach to Bio-Molecular Event Extraction*
Sebastian Riedel, Hong-Woo Chun, Toshihisa Takagi and Jun'ichi Tsujii

11:20–11:40 *High-precision biological event extraction with a concept recognizer*
K. Bretonnel Cohen, Karin Verspoor, Helen Johnson, Chris Roeder, Philip Ogren, William Baumgartner, Elizabeth White and Lawrence Hunter

11:40–11:55 *A memory-based learning approach to event extraction in biomedical texts*
Roser Morante, Vincent Van Asch and Walter Daelemans

11:55–12:10 *Extraction of biomedical events using case-based reasoning*
Mariana Neves, José-María Carazo and Alberto Pascual-Montano

12:10–12:25 *Biomedical Event Annotation with CRFs and Precision Grammars*
Andrew MacKinlay, David Martinez and Timothy Baldwin

Friday, June 5, 2009 (continued)

12:30–14:00 Lunch

Session 3: Poster presentations

14:00–14:15 Introduction to poster presentations

14:15–15:30 Poster presentations

Molecular event extraction from Link Grammar parse trees
Jörg Hakenberg, Illes Solt, Domonkos Tikk, Luis Tari, Astrid Rheinländer, Nguyen Quang Long, Graciela Gonzalez and Ulf Leser

Tunable Domain-Independent Event Extraction in the MIRA Framework
Georgi Georgiev, Kuzman Ganchev, Vassil Momchev, Deyan Peychev, Preslav Nakov and Angus Roberts

BioEve: Bio-Molecular Event Extraction from Text Using Semantic Classification and Dependency Parsing
Syed Toufeeq Ahmed, Radhika Nair, Chintan Patel and Hasan Davulcu

From Protein-Protein Interaction to Molecular Event Extraction
Rune Sætre, Makoto Miwa, Kazuhiro Yoshida and Jun'ichi Tsujii

A Multi-Phase Approach to Biomedical Event Extraction
Hyoung-Gyu Lee, Han-Cheol Cho, Min-Jeong Kim, Joo-Young Lee, Gumwon Hong and Hae-Chang Rim

Supervised Classification for Extracting Biomedical Events
Arzucan Ozgur and Dragomir Radev

Biomedical Event Detection using Rules, Conditional Random Fields and Parse Tree Distances
Farzaneh Sarafraz, James Eales, Reza Mohammadi, Jonathan Dickerson, David Robertson and Goran Nenadic

15:30–16:00 Coffee break

Friday, June 5, 2009 (continued)

Session 4: Oral presentations and Discussion

16:00–16:20 *Syntactic Dependency Based Heuristics for Biological Event Extraction*
Halil Kilicoglu and Sabine Bergler

16:20–16:35 *Analyzing text in search of bio-molecular events: a high-precision machine learning framework*
Sofie Van Landeghem, Yvan Saeys, Bernard De Baets and Yves Van de Peer

16:35–16:50 *Exploring ways beyond the simple supervised learning approach for biological event extraction*
György Móra, Richárd Farkas, György Szarvas and Zsolt Molnár

16:50–17:30 Discussion