##

 CALL FOR PAPERS

>> LREC2008 - Towards a Shared Task for Multiword Expressions (MWE 2008)
<<

 endorsed by the ACL Special Interest Group on the Lexicon (SIGLEX)

Date: Sunday, 1 June 2008
Location: Marrakech, Morocco

Workshop web page: http://multiword.sf.net/mwe2008/

##

In recent years, considerable progress has been made in our
understanding of
multiword expressions (MWE), the development of algorithms for their
automatic
extraction from corpora, and the automatic identification of additional
properties such as morphosyntactic preferences or the interpretation of
semi-compositional expressions.

It is difficult to compare results of the many published studies on MWEs
and
obtain a broader perspective, though, because algorithms and implemented
systems have been evaluated on vastly different gold standards and
corpora, in
different languages, for different subtypes of MWEs, etc. In order to
make the
next big step forward, the field of MWE research needs a shared task in
which
different approaches are applied to the same data sets, allowing
completely
new insights to be gained. Since there is as yet not a clear and
universally
accepted definition of multiword expressions, the first instalment of
this
shared task will be of a more exploratory nature than the competitions
that
have been carried out in other areas of computational linguistics.

The MWE 2008 workshop is primarily intended as a forum for collecting,
sharing
and exploiting MWE evaluation resources. We have solicited contributions
of
such resources in a separate call.

After collection of the resources, teams are invited to participate in a
shared task by evaluating their MWE extraction algorithms on data sets
downloaded from http://multiword.sf.net/. Further instructions will be
made
available on the workshop web site, and the full collection of data sets
will
be online by February 1, 2008 at the latest.

There will be three different types of submissions to the workshop:

(1) short papers describing data sets and other evaluation resources

made
      freely available on the community Web page;

(2) shared task participants, who evaluate an algorithm or MWE
extraction
      system on multiple data sets and discuss implications of their
results;

(3) regular papers addressing the evaluation and comparison of multiword
      extraction algorithms (but not limited to these topics).

With this call, we invite submissions of regular papers, in particular
(but not limited to) research on:

(a) Linguistic analysis of MWEs based on language resources (such as
corpora)
and the impact that these studies have on NLP applications. We
particularly
welcome papers that perform a cross-linguistic analysis of MWEs,
identify
variation across languages, text types, domains, etc. or investigate the
variability of MWEs.

(b) Typologies of MWEs: Papers that describe classes of MWEs and their
representation in language resources, discuss different approaches to
the
definition and classification of MWEs, or apply new MWE typologies to
the
evaluation of computational techniques.

(c) The evaluation and comparison of multiword extraction techniques: Do
methods generalise across languages, text types, different classes of
MWEs,
etc.? How useful and essential is linguistic knowledge and a theoretical
understanding of MWEs? Is fully automatic extraction feasible or will
manual
intervention always be necessary?

(d) Evaluation methodology and the creation of gold standards for MWEs.
Papers
should address theoretical and technical issues, while descriptions of
existing resources may be submitted as short papers for the shared task.
Topics of particular interest are novel types of gold standards (such as
human
ratings from expert and non-expert subjects, language resources derived
from
the Web, etc.), inter-annotator agreement in the manual validation of
candidate lists (which is often fairly low) and the task-based
evaluation of
MWE resources in NLP applications.

SUBMISSION INFORMATION

Regular papers must adhere to the format of LREC proceedings (preferably
using the style files provided
on the conference Web site) and must not exceed eight (8) pages,
including references.
Only submissions in PDF format will be considered.

Short papers describing evaluation resources and shared task

participants will
be allowed four (4) pages, using the same formatting.

The papers must be submitted no later than 23:59 GMT February 29, 2008.
Papers submitted after that time cannot be reviewed.

Further submission instructions will be provided on the workshop Web
site.

IMPORTANT DATES

Paper submission deadline: February 29, 2008
Notification of acceptance: March 28, 2008
Camera ready papers due: April 4, 2008
Workshop date: June 1, 2008

PROGRAM COMITTEE

Iñaki Alegria   University of the Basque Country (Spain)
Timothy Baldwin   Stanford University (USA); U of Melbourne
(Australia)
Colin Bannard   Max Planck Institute (Germany)
Francis Bond   NTT Communication Science Laboratories (Japan)
Gaël Dias   Beira Interior University (Portugal)
Ulrich Heid   Stuttgart University (Germany)
Kyo Kageura   University of Tokyo (Japan)
Rosamund Moon   University of Birmingham (UK)
Diana McCarthy   University of Sussex (UK)
Eric Laporte   University of Marne-la-Vallee (France)
Preslov Nakov   University of California, Berkeley (USA)
Jan Odijk   University of Utrecht (The Netherlands)
Stephan Oepen   Stanford University (USA); U of Oslo (Norway)
Darren Pearce   University of Sussex (UK)
Pavel Pecina   Charles University (Czech Republic)
Scott Piao   University of Manchester (UK)
Violeta Seretan   University of Geneva (Switzerland)
Suzanne Stevenson   University of Toronto (Canada)
Beata Trawinski   University of Tuebingen (Germany)
Kiyoko Uchiyama   Keio University (Japan)
Begoña Villada Moirón University of Groningen (The Netherlands)
Aline Villavicencio   Federal University of Rio Grande do Sul (Brazil)

WORKSHOP CHAIRS

Nicole Grégoire
University of Utrecht, The Netherlands

Stefan Evert
University of Osnabrueck, Germany

Brigitte Krenn
Austrian Research Institute for Artificial Intelligence (ÖFAI), Austria

CONTACT

For any inquiries regarding the workshop please contact Nicole Grégoire
(Nicole.Gregoire@let.uu.nl).