Australasian Language Technology Association Workshop 2011

(ALTA 2011)

Canberra, Australia
1 – 2 December 2011

Editors:

Diego Molla
David Martinez

ISBN: 978-1-62276-510-2
ISSN: 1834-7037

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2011) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2012)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Contents

Non-reviewed papers **1**

Discovery in Text: Visualisation, Topics and Statistics
Wray Buntine **2**

OzCLO: The Australian Computational Linguistic Olympiad
Dominique Estival **3**

Automatic Grading of Evidence: the 2011 ALTA Shared Task
Diego Mollá and Abeed Sarker **4**

Peer-reviewed papers: Oral presentations **9**

A Particle Filter algorithm for Bayesian Wordsegmentation
Benjamin Börschinger and Mark Johnson **10**

Formalizing Semantic Parsing with Tree Transducers
Bevan Jones, Mark Johnson and Sharon Goldwater **19**

Parsing in Parallel on Multiple Cores and GPUs
Mark Johnson **29**

Using Language Models and Latent Semantic Analysis to Characterise the N400m Neural Response
Mehdi Parviz, Mark Johnson, Blake Johnson and Jon Brock **38**

A Forensic Authorship Classification in SMS Messages: A Likelihood Ratio Based Approach Using N-gram
Shunichi Ishihara **47**

Classifying Domain-Specific Terms Using a Dictionary
Su Nam Kim and Lawrence Cavedon **57**

Frontier Pruning for Shift-Reduce CCG Parsing
Stephen Merity and James R. Curran **66**

Predicting Thread Linking Structure by Lexical Chaining
Li Wang, Diana Mccarthy and Timothy Baldwin **76**

Development of a Corpus for Evidence Based Medicine Summarisation
Diego Mollá and Maria Elena Santiago-Martinez **86**

Collocations in Multilingual Natural Language Generation: Lexical Functions meet Lexical Functional Grammar
François Lareau, Mark Dras, Benjamin Börschinger and Robert Dale **95**

Outcome Polarity Identification of Medical Papers
Abeed Sarker, Diego Mollá and Cécile Paris **105**

Topic Modeling for Native Language Identification
Sze-Meng Jojo Wong, Mark Dras and Mark Johnson **115**

Peer-reviewed papers: Poster presentations **125**

A word-based approach for diacritic restoration in Māori
John Cocks and Te Taka Keegan **126**

The Interpretation of Complement Anaphorae: the case of The Others
Nobuagi Akagi and Francesco-Alessio Ursini **131**

The Interpretation of Plural Pronouns in Discourse: The Case of They
Francesco-Alessio Ursini and Nobuagi Akagi **140**

Learning from student responses: A domain-independent natural language tutor
Jenny Mcdonald, Alistair Knott, Richard Zeng and Ayelet Cohen **148**

Detection of child exploiting chats from a mixed chat dataset as a text classification task
Md. Waliur Rahman Miah, John Yearwood and Sid Kulkarni **157**

ENGAGE: Automated Gestures for Animated Characters
Marcin Nowina-Krowicki, Andrew Zschorn, Michael Pilling and Steven Wark **166**