

First SIGHAN Workshop on Chinese Language Processing 2002

Held at COLING 2002

**Taipei, Taiwan
1 September 2002**

ISBN: 978-1-62748-370-4

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© (2002) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2013)

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

TABLE OF CONTENTS

Extraction of Translation Unit from Chinese-English Parallel Corpora	1
<i>Baobao Chang, Pernilla Danielsson, Wolfgang Teubert</i>	
Some Considerations on Guidelines for Bilingual Alignment and Terminology Extraction	6
<i>Lawrence Cheung, Tom Lai, Robert Luk, Oi Yee Kwong, King Kui Sin, Benjamin K. Tsou</i>	
Developing Guidelines for the Annotation of Anaphors in the Chinese Treebank	11
<i>Susan Converse</i>	
Finding the Better Indexing Units for Chinese Information Retrieval	18
<i>Hongzhao He, Pilian He, Jianfeng Gao, Changning Huang</i>	
Categorical Ambiguity and Information Content - A Corpus-based Study of Chinese.....	25
<i>Chu-Ren Huang, Ru-Yng Chang</i>	
PCFG Parsing for Restricted Classical Chinese Texts	30
<i>Liang Huang, Yinan Peng, Huan Wang, Zhenyu Wu</i>	
A Knowledge Based Approach to Identification of Serial Verb Construction in Chinese-to-Korean Machine Translation System	36
<i>Dong-il Kim, Zheng-Cui Jinji-Li, Jong-Hyeok Lee</i>	
Learning Case-based Knowledge for Disambiguating Chinese Word Segmentation: A Preliminary Study	43
<i>Chunyu Kit, Haihua Pan, Hongbiao Chen</i>	
Corpus-Based Pinyin Name Resolution	50
<i>Kui-Lam Kwok, Peter Deng</i>	
An Indexing Method Based on Sentences	57
<i>Li Li, Chunfa Yuan, K. F. Wong, Wenjie Li</i>	
Design of Chinese Morphological Analyzer.....	62
<i>Huihsin Tseng, Keh-Jiann Chen</i>	
A Word Segmentation Method With Dynamic Adapting To Text Using Inductive Learning	69
<i>Zhongjian Wang, Kenji Araki, Koji Tochinal</i>	
Using the Segmentation Corpus to Define an Inventory of Concatenative Units for Cantonese Speech Synthesis	74
<i>Wai Yi Peggy Wong, Shui-duen Chan, Chris Brew, Mary E. Beckman</i>	
Extracting Pronunciation-translated Names from Chinese Texts using Bootstrapping Approach	79
<i>Jing Xiao, Jimin Liu, Tat-Seng Chua</i>	
Combining Classifiers for Chinese Word Segmentation	85
<i>Nianwen Xue, Susan P. Converse</i>	
WSD and Closed Semantic Constraint	92
<i>Jiangsheng Yu</i>	
Automatic Recognition of Chinese Unknown Words¹ Based on Roles Tagging.....	97
<i>Kevin Zhang, Qun Liu, Hao Zhang, Xue-Qi Cheng</i>	
Chinese Base-Phrases Chunking	104
<i>Yuqi Zhang, Qiang Zhou</i>	
Learning Rules for Chinese Prosodic Phrase Prediction.....	109
<i>Sheng Zhao, Jianhua Tao, Lianhong Cai</i>	
A Knowledge-based Approach to Text Classification.....	116
<i>Jingbo Zhu, Tianshun Yao</i>	
Author Index	