Proceedings of the

Second Chinese Language Processing Workshop

Held in conjunction with
The 38th Annual Meeting of the
Association for Computational Linguistics

Edited by

Martha Palmer
Mitch Marcus
Aravind Joshi
Fei Xia

8 October 2000
Hong Kong University of Science and Technology (HKUST)
Hong Kong

Proceedings of the

Second Chinese Language Processing Workshop

Held in conjunction with
The 38th Annual Meeting of the
Association for Computational Linguistics

Edited by

Martha Palmer
Mitch Marcus
Aravind Joshi
Fei Xia

8 October 2000
Hong Kong University of Science and Technology (HKUST)
Hong Kong

Order copies of this and other ACL workshop proceedings from:

Association for Computational Linguistics (ACL)
75 Paterson Street, Suite 9
New Brunswick, NJ 08901
USA
Tel: +1-732-342-9100
Fax: +1-732-342-9339
acl@aclweb.org

SPONSORS:

SIGDAT
SIGLEX
SIGPARSE

INDUSTRY SPONSOR:

Intel China Research Center

INVITED SPEAKER:

Chin-Chuan Cheng (City University of Hong Kong)

ORGANIZERS:

Martha Palmer
Mitch Marcus
Aravind Joshi
Fei Xia

PROGRAM COMMITTEE:

Srinivas Bangalore	(AT&T Research Labs)
Keh-Jiann Chen	(Academia Sinica)
Zhengdong Dong	(Hownet)
Shengli Feng	(Univ. of Kansas)
Kok Wee Gan	(HKUST)
Laurie Gerber	(Univ. of Southern CA/ISI)
Changning Huang	(Microsoft Research China)
Chu-Ren Huang	(Academia Sinica)
Wanying Jin	(New Mexicon State Univ.)
Kim-Teng Lua	(National Univ. of Singapore)
John Kovarik	(Department of Defense)
K.L. Kwok	(Queens College)
Mary Ellen Okurowski	(Department of Defense)
Fuji Ren	(Hiroshima City Univ.)
Richard Sproat	(AT&T Research Lab)
Keh-Yih Su	(Behavior Design Corporation)
Maosong Sun	(Tsinghua Univ.)
Chew Lim Tan	(National Univ. of Singapore)
Benjamin K Tsou	(City Univ. of Hong Kong)
Amy Weinberg	(Univ. of Maryland)
Ralph Weischedel	(BBN)
Andi Wu	(Microsoft)
Dekai Wu	(HKUST)
Nianwen Xue	(Univ. of Delaware)
Jin Yang	(Systran)
Shiwen Yu	(Peking Univ.)
Chunfa Yuan	(Tsinghua Univ.)
Joe Zhou	(Intel China Research Center)
Qiang Zhou	(Tsinghua Univ.)

FURTHER INFORMATION:

Martha Palmer, Mitch Marcus, Aravind Joshi, and Fei Xia
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
Email: {mpalmer,mitch,joshi,fxia}@linc.cis.upenn.edu

ACKNOWLEDGMENT:

Special thanks to Yen-Lin Yin at the Univ. of Pennsylvania for helping us to organize this workshop.

WORKSHOP PROGRAM

Sunday, 8 October 2000

8:45-9:00 Welcome

9:00-9:50 *Invited Talk: Zero Anaphors in Chinese Discourse Processing*
Chin-Chuan Cheng

9:50-10:10 *Sense-Tagging Chinese Corpus*
Hsin-Hsi Chen and Chi-Ching Lin

10:10-10:30 *Enhancement of a Chinese Discourse Marker Tagger with C4.5*
Benjamin K. T'sou, Tom B.Y. Lai, Samuel W.K. Chan, Weijun Gao and Xuegang Zhan

10:30-11:00 *Break*

11:00-11:20 *Two Statistical Parsing Models Applied to the Chinese Treebank*
Daniel M. Bikel and David Chiang

11:20-11:40 *Using Co-occurence Statistics as an Information Source for Partial Parsing of Chinese*
Elliott Franco Drabek and Qiang Zhou

11:40-12:00 *Statistics Based Hybrid Approach to Chinese Base Phrase Identification*
Tie-jun Zhao, Mu-yun Yang, Fang Liu, Jian-min Yao and Hao Yu

12:00-12:20 *A Block-Based Robust Dependency Parser for Unrestricted Chinese Text*
Ming Zhou

12:20-1:30 *Lunch*

1:30-2:30 *Poster session*

2:30-2:50 *Knowledge Extraction for Identification of Chinese Organization Names*
Keh-Jiann Chen and Chao-jan Chen

2:50-3:10 *Statistically-Enhanced New Word Identification in a Rule-Based Chinese System*
Andi Wu and Zixin Jiang

3:10-3:30 *A Trainable Method for Extracting Chinese Entity Names and Their Relations*
Yimin Zhang and Joe F Zhou

3:30-4:00 *Break*

4:00-4:20 *Sinica Treebank: Design Criteria, Annotation Guidelines, and On-line Interface*
Chu-Ren Huang, Feng-Yi Chen, Keh-Jiann Chen, Zhao-ming Gao and Kuang-Yu Chen

4:20-4:40 *Comparing Lexicalized Treebank Grammars Extracted from Chinese, Korean, and English*
Fei Xia, Chunghye Han, Martha Palmer and Aravind Joshi

4:40-5:00 *The Research of Word Sense Disambiguation Method Based on Co-occurrence Frequency of Hownet*
Erhong Yang, Guoqing Zhang and Yongkui Zhang

5:00-6:00 *Panel: Prioritizing Chinese Language Processing Resources*
Keh-Jiann Chen, Chu-Ren Huang, Bing Swen, Benjamin K. T'sou,
Joe Zhou and Ming Zhou

Table of Contents