

计算语言学教育部重点实验室 2012 成果清单

一、论文

Dongxu Han, Baobao Chang. Approaches to domain adaptive Chinese segmentation model. In proceedings of the 2012 International Conference on Natural Language Processing and Knowledge Engineering. 2012.

Mairgup Mansur, Baobao Chang. Using Web-based re-ranking in Two-step Conditional Random Field Model for Machine Transliteration. In proceedings of the 2012 International Conference on Natural Language Processing and Knowledge Engineering. 2012

董秀芳. Lexicalization in the history of the Chinese language, in Xing, Janet. (ed.) 2012. Newest Trends in the Study of Grammaticalization and Lexicalization in Chinese. Berlin & Boston: De Gruyter Mouton: 235-274.

董秀芳.上古汉语议论语篇的结构与特点：兼论联系语篇结构分析虚词的功能，《中国语文》2012 年第 4 期：356-366 页。

董秀芳.《语义关系与词库》导读, M. Lynne Murphy 著《Semantic Relations and Lexicon (语义关系与词库)》, 世界图书出版公司: 11-35 页, 2012 年 8 月。

Liwei Chen, Yansong Feng*, Lei Zou, Dongyan Zhao: Explore Person Specific Evidence in Web Person Name Disambiguation. EMNLP-CoNLL 2012: 832-842 (EI: 20133716718381)

Yichen Jiang, Aixia Jia, Yansong Feng*, Dongyan Zhao: Recommending academic papers via users' reading purposes. ACM Conference on Recommender System 2012: 241-244 (EI: 20124215578149) (short paper)

Jie Han, Sujian Li, Zhifang Sui. Experiment Research on Automated Essay Scoring for Nonnative Chinese Speakers. ICIC Express Letters, Part B: Applications, v 3, n 5, p 1285-1291, 2012. (EI)

Xun Wang, Lei Wang, Jiwei Li, Sujian Li. Exploring Simultaneous Keyword and Key Sentence Extraction: Improve Graph-based Ranking Using Wikipedia. CIKM 2012 poster. pp. 2619-2622, Oct. 2012, Sheraton, Maui Hawaii.

Xinfan Meng, Furu Wei, Xiaohua Liu, Ming Zhou, Sujian Li, Houfeng Wang: Entity-centric topic-oriented opinion summarization in twitter. KDD 2012: 379-387.

Yang Song, Jing Jiang, Wayne Xin Zhao, Sujian Li, Houfeng Wang. Joint Learning for Coreference Resolution with Markov Logic. EMNLP 2012, pages 1245–1254, Jeju Island, Korea.

Jiwei Li, Sujian Li, Xun Wang, Ye Tian, Baobao Chang, Update Summarization Using a Multi-level Hierarchical Dirichlet Process Model, Coling 2012, pages 1603-1618, Dec. 2012, Mumbai, India.

Xun Wang, Sujian Li, Jiwei Li, Wenjie Li, Implicit Discourse Relation Recognition by Selecting Typical Training Examples, Coling 2012, pages 2757-2772, Dec. 2012, Mumbai, India.

Longkai Zhang, Sujian Li, Houfeng Wang, Ni Sun, and Xinfan Meng, Constructing Chinese Abbreviation Dictionary: A Stacked Approach, Coling 2012, pages 3055-3070, Dec. 2012, Mumbai, India.

刘扬, 陆顾婧. 汉英双语概念对应的实证研究, 云南师范大学学报 (哲学社会科学版), 2012, 44(1).

Zhifang Sui, Wei Kang, Ye Tian, Synchronously Extracting Instances and Attributes for the Concepts from the Web, International Journal of Knowledge and Language Processing, Vol.3 No. 3, 2012

Wenjie Li, Jiwei Li, Ye Tian and Zhifang Sui, Fine-Grained Classification of Named Entities by Fusing Multi-features, In: Proceedings of The 24th International Conference on Computational Linguistics (COLING-2012), Mumbai, India, 2012

Duan Huiming, Zhifang Sui, Ye Tian, Wenjie Li, The CIPS-SIGHAN CLP 2012 Chinese Word Segmentation on MicroBlog Corpora Bakeoff, The Second CIPS-SIGHAN Joint Conference on Chinese Language Processing, 2012

王鑫、穗志方, 基于依存树距离的语义角色识别方法, 《中文信息学报》, 2012 年第 2 期

李文杰、穗志方, 基于并列结构的概念实例和属性的同步提取方法, 《中文信息学报》, 2012 年第 2 期

Weiwei Sun and Hans Uszkoreit. 2012. Capturing paradigmatic and syntagmatic lexical relations: Towards accurate Chinese part-of-speech tagging. ACL.

Weiwei Sun and Xiaojun Wan. 2012. Reducing approximation and estimation errors for Chinese lexical processing with heterogeneous annotations. ACL.

Minwei Feng, Weiwei Sun and Hermann Ney. 2012. Semantic Cohesion Model for Phrase-based SMT. COLING.

Xu Sun, Houfeng Wang, Wenjie Li. Fast Online Training with Frequency-Adaptive Learning Rates for Chinese Word Segmentation and New Word Detection. In ACL. 253–262. 2012.

Liqiang Guo, Xiaojun Wan. Exploiting Syntactic and Semantic Relationships between Terms for Opinion Retrieval. Journal of the American Society for Information Science and Technology (JASIST), 63(11):2269-2282.

Liqliang Guo, Xiaojun Wan. S2ORM: Exploiting Syntactic and Semantic Information for Opinion Retrieval. In WWW2012, pages 517-518. (Poster Paper)

Weiwei Sun, Xiaojun Wan. Reducing approximation and estimation errors for Chinese lexical processing with heterogeneous annotations. InACL2012, pages 232-241. (Long Paper)

Rui Yan, Xiaojun Wan, Mirella Lapata, Pu-Jen Cheng, Xiaoming Li. Visualizing Timelines: Evolutionary Summarization via Iterative Reinforcement between Text and Image Streams. In CIKM2012, pages 275-284. (Full Paper, Acceptance Rate = 13.4%)

Rui Yan, Zi Yuan, Xiaojun Wan, Yan Zhang, Xiaoming Li. Hierarchical Graph Summarization: Leveraging Hybrid Information through Visible and Invisible Linkage. In PAKDD2012, pages 97-108. (Full Paper)

Xiaojiang Huang, Xiaojun Wan, Jianguo Xiao. BiCWS: Mining Cognition Differences from Bilingual Web Search Results. In WISE2012, pages 58-71. (Regular Paper)

Xiaojiang Huang, Xiaojun Wan, Jianguo Xiao. Learning to Find Comparable Entities on the Web. In WISE2012, pages 16-29. (Regular Paper)

Xiaojun Wan. A Comparative Study of Cross-Lingual Sentiment Classification. In WI2012, 8 pages. (Regular Paper)

Xinjie Zhou, Xiaojun Wan, Jianguo Xiao. Cross-Language Opinion Target Extraction in Review Texts. In ICDM2012, pages 1200-1205. (Short Oral Paper, Acceptance Rate [Regular + Short] =20%)

Xiaojun Wan. Update Summarization Based on Co-Ranking with Constraints. In COLING2012, pages 1291-1300. (Poster Paper)

Xinfan Meng, Furu Wei, Xiaohua Liu, Ming Zhou, Ge Xu, Houfeng Wang Cross-Lingual Mixture Model for Sentiment Classification. Proceedings of the conference of the 50th Annual Meeting of the association for Computational Linguistics (ACL 2012)

Xu Sun, Houfeng Wang, Wenjie Li. Fast Online Training with Frequency-Adaptive Learning Rates for Chinese Word Segmentation and NewWord Detection. Proceedings of the conference of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012)

Yang Song, Jing Jiang, Xin Zhao and Houfeng Wang. Joint Learning for Coreference Resolution with Markov Logic. Proceedings of the conference on Empirical Methods in Natural Language Processing and Natural Language Learning (EMNLP-CoNLL 2012)

Xinfan Meng, Furu Wei, Xiaohua Liu, Ming Zhou, Houfeng Wang. Entity-Centric Topic-Oriented Opinion Summarization in Twitter. Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2012).

Xu Sun, Naoaki Okazaki, Jun'ichi Tsujii, Houfeng Wang. Learning Abbreviations from Chinese and English Terms by Modeling Non-local Information. ACM Trans. on Asian Language Information Processing (accepted, to appear)

Ge Xu, Chunren Huang, Houfeng huang. Extracting Chinese Product Features:Representing a Sequence by a Set of Skip-Bigrams. Proceedings of the 13th Chinese Lexical Semantics Workshop(CLSW2012) (to appear in LNAI)

Longkai Zhang, Sujian Li, Houfeng Wang and Ni Sun . Constructing Chinese Abbreviation Dictionary: A Stacked Approach. Proceedings of the 24th International Conference on Computational Linguistics(COLING 2012)

Zhengyan He and Houfeng Wang. A Comparison and Improvement of Online Learning Algorithms for Sequence Labeling. Proceedings of the 24th International Conference on Computational Linguistics(COLING 2012)

Xinfan Meng, Furu Wei, Ge Xu, Xiaohua Liu, Ming Zhou and Houfeng Wang. Lost in Translations? Building Sentiment Lexicons Using Context Based Machine Translation. Proceedings of the 24th International Conference on Computational Linguistics(COLING 2012)

焦妍, 王厚峰, 张龙凯; 基于条件随机场与 Web 数据的缩略语预测[J]中文信息学报.2012(02)

张龙凯, 王厚峰, 文本摘要问题中的句子抽取方法研究[J]中文信息学报.2012(02)

Fei Wang, Yunfang Wu and Likun Qiu. 2012. Exploiting discourse relations for sentiment analysis. *Proceedings of COLING-2012*, Pp 1311-1320.

Fei Wang, Yunfang Wu. 2012. Mining market trend from blog titles based on lexical semantic similarity. *Proceedings of 12th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2012)*, Pp 261-273.

Peng Jin, John Carroll, Yunfang Wu and Diana McCarthy. 2012. Distributional similarity for Chinese: exploiting characters and radicals. *Mathematical Problems in Engineering*.(SCI Index).

Peng Jin and Yunfang Wu. 2012. SemEval-2012 task 4: evaluating Chinese word similarity. *Proceedings of First Joint Conference on Lexical and Computational Semantics (SEM)*, Pp 374-377.

晋陕甘宁部分方言古全浊声母的今读 《咸阳师范学院学报》2012 年第 5 期

方言地理、方言分区和谱系分类 《龙岩学院学报》2012 年第 4 期

汉语词汇地理学的一般程序——以“冰雹”的研究为例 《地理科学研究》(OA 期刊)2012 年第 2 期

袁毓林。《动词内隐性否定的语义层次和溢出条件》，《中国语文》2012 年第 3 期，第 99-113 页。中国人民大学复印报刊资料《语言文字学》2012 年第 6 期，第 16-30 页。全文转载。

袁毓林。《修辞学家可以向邻近学科学些什么？》，《当代语言学》2012 年第 3 期，第 286-300 页。

魏雪、袁毓林。To Construct the Interpretation Templates for the Chinese Noun Compounds Based on Semantic Classes and Qualia Structures, in *Papers of the 26th Pacific Asia Conference on Language, Information and Computation (PACLIC26)*, 8-10 November, 2012, Bali, Indonesia, published by Faculty of Computer Science, Universitas of Indonesia, pp. 609-619.

袁毓林。《由述结式构成的“把”字句的语义类型》，澳门大学《南国人文学刊》2012 年第 2 期，第 118-128 页。

袁毓林、王明华。《隐性否定溢出的例证调查和博弈论分析》，复旦大学汉语言文字学科《语言研究集刊》（第九辑），第 97-112 页，上海辞书出版社，2012 年 10 月。

Lei Zou, Lei Chen, M. Tamer Özsu, Dongyan Zhao, “Answering Pattern Match Queries in Large Graph Databases via Graph Embedding”, The International Journal on Very Large Data Bases, The VLDB Journal, Vol. 21, No.1 pp. 97-120, 2012.

Tinghua Wang, Dongyan Zhao, Shengfeng Tian, An Overview of Kernel Alignment and Its Applications, Artificial Intelligence Review, DOI:10.1007/s10462-012-9369-4, 2012.

Liwei Chen, Yansong Feng, Yidong Chen, Lei Zou and Dongyan Zhao, “Towards Automatically Constructing Knowledge Bases from Chinese Resources”, Annual Meeting of the Association for Computational Linguistics Student Research Workshop (ACL-SRW), Jeju Island, Korea, pp. 67-72, Jul. 2012.

Liwei Chen, Yansong Feng, Lei Zou and Dongyan Zhao, “Explore Person Specific Evidence in Web Person Name Disambiguation”, Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Jeju Island, Korea, pp. 832-842, Jul. 2012.

Yichen Jiang, Aixia Jia, Yansong Feng, Dongyan Zhao, “Recommending Academic Papers via Users’ Reading Purposes”,
The ACM Conference on Recommender Systems (ACM RecSys), Dublin, Ireland, pp. 241-244, Sep. 2012.

Wei Wang, Dongyan Zhao, “Ontology-based Event Modeling for Semantic Understanding of Chinese News Story”, CCF
Conference on Natural Language Processing & Chinese Computing (NLP&CC 2012), Beijing, China, pp. 58-68, Nov.
2012.

Wei Wang, Dongyan Zhao, “Chinese News Event 5W1H Semantic Elements Extraction for Event Ontology Population”.
World Wide Web Conference PhD Symposium (WWW PhD Symposium), Lyon, France, pp. 197-202, Apr. 2012.

The Conference on Empirical Methods on Natural Language Processing, EMNLP, 2012.

Lei Zou, Lei Chen, M. Tamer Özsu, Dongyan Zhao. Answering pattern match queries in large graph databases via graph
embedding, VLDB Journal, VLDB J, 2012.

Lei Wang, Shiwen Yu, Xuefeng Zhu, Yun Li. Chinese Idiom Knowledge Base for Chinese Information Processing, CLSW
2012, Wuhan, China, July 6-8, 2012, pp 302-310.

Xiaojun Zou, Hua Zhang, Junfeng Hu*. Internet Based Industry Community Discovery And Its Application To Industry Survey, Proc. the Internatioal Conference on Automatic Control and Artificial Intelligence (ACAI 2012). p1255-1259, 2012.

胡俊峰, 邹小军, 陈玉忠, 任冠华, 基于互联网产业社区分析的标准化体系研究, 中国标准化, 第三期, P83-88, 2012.

二、专著

《汉语句子的焦点结构和语义解释》，商务印书馆，2012年，第1-362页，袁毓林。

三、科研项目

类型	批准部门	编号	负责人	时间	经费
----	------	----	-----	----	----

国家社科基金后期资助项目	国家社科基金	11FYY013	吴云芳	2012.01- 2013.06	15 万
	南京金娃娃软件科技有限公 司		吴云芳	2012.01- 2015.12.	50 万
国家社科基金重大项目“基于 多学科视域的认知研究”（北 大哲学系周北海主持）	全国哲学社会科学规划办	12&ZD119	董秀芳	2012.10- 2017.10	
国家 863 计划	国家科技部	2012AA011101	王厚峰	2012.1- 2014.12	
863 计划	科技部	2012AA011101 大规 模汉语语义基础资 源库和知识库设计 构建及工具平台	王厚峰	2012.1- 2014.12	80 万

国家自然科学基金	国家自然科学基金委	61170166	万小军	2012.1- 2015.12	54 万
国家社科重大项目	国家社科规划办	12&ZD119	周北海	2012.7- 2017.6	80 万 元
国家自然科学基金青年	国家自然科学基金委员会	61202233	冯岩松	2012/01- 2015/12	25 万
国家自然科学基金面上	国家自然科学基金委员会	61272344	赵东岩	2012/01- 2016/12	20 万
国家高技术发展研究计划 (863 项目)	国家科技部	2012AA011101	王厚峰 (首席)	2012-2014	659 万 元
国家社科基金重大招标项目	国家社科基金委员会	12&ZD175	袁毓林	2012 年 10 月-2017 年 10 月	80+70 =150 万元

面上项目	国家自然科学基金委	31170972	周晓林	2012年1月 -2015年12月	70 万 元
国家社会科学基金后期资助项目	国家社会科学基金委员会	12FYY005	袁毓林	2012年8月 -2014年6月	15 万 元